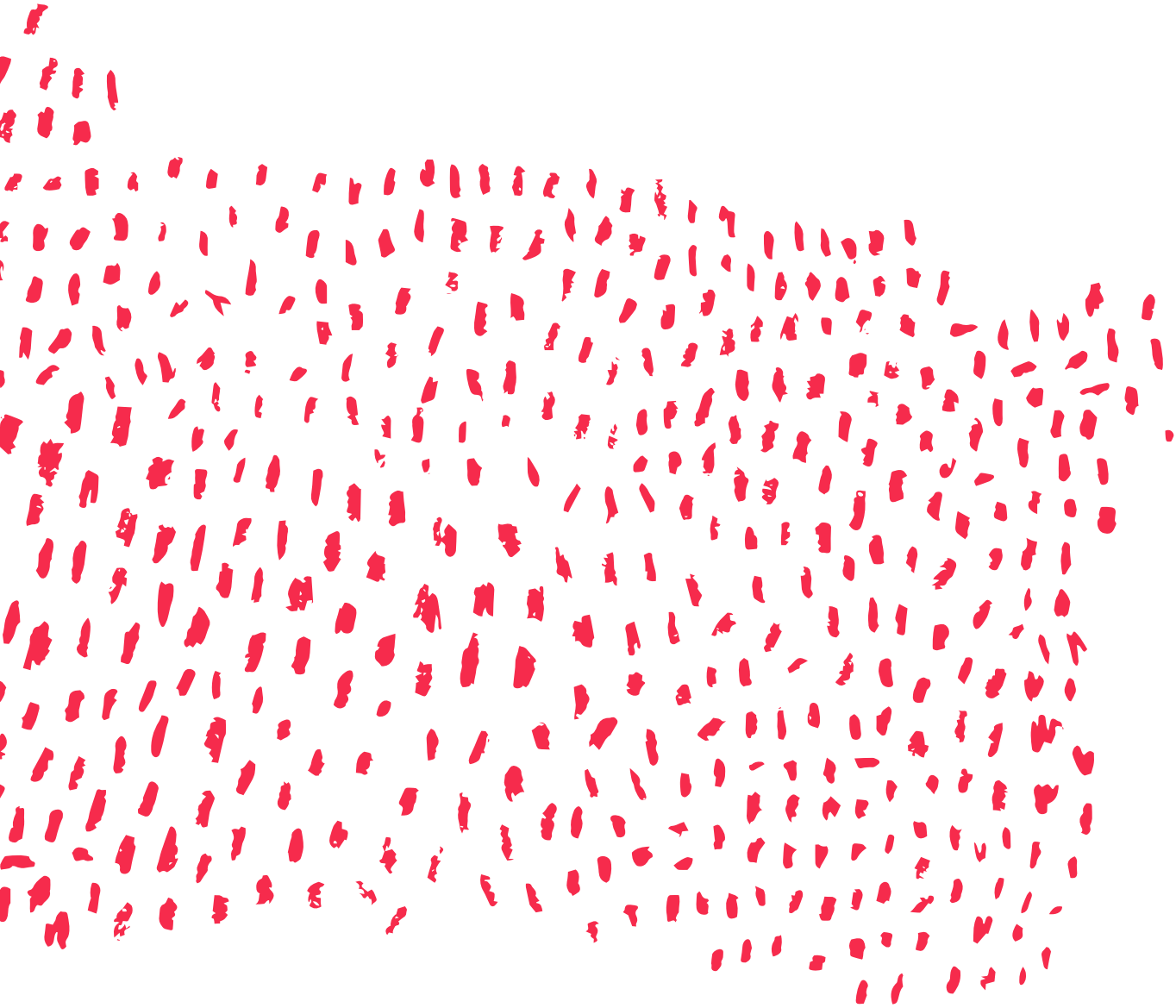
DATNAV

How to navigate digital data for human rights research

THE ENGINE ROOM

benetech
TECHNOLOGY SERVING HUMANITY

AMNESTY INTERNATIONAL



Eyewitness Media Hub

University of Essex

GULF CENTER FOR HUMAN RIGHTS

HRDAG

Syria Justice and Accountability Centre

WITNESS SEE IT FILM IT CHANGE IT

Contributors

Our thanks go to the following **writing sprint contributors:**

- Allison Corkery, Center for Economic and Social Rights
- Sam Dubberley, Eyewitness Media Hub and Human Rights and Big Data Project, University of Essex
- Scott Edwards, Amnesty International
- Lisa Gutermuth
- Danna Ingleton, Amnesty International
- Christoph Koettl, Amnesty International
- Jule Krüger, Human Rights Data Analysis Group (HRDAG)
- Chris Michael, Collaborations for Change
- Ella McPherson, Department of Sociology and the Centre of Governance and Human Rights, University of Cambridge
- Shabnam Mojtahedi, Syrian Justice and Accountability Center
- Chitra Nagarajan
- Zara Rahman, The Engine Room
- Elsa Saade, Gulf Center for Human Rights
- Collin Sullivan, Benetech
- Jackie Zammuto, WITNESS

Our thanks also go to the following **people who contributed to the project as reviewers** of this document, community call participants, or interviewees in the earlier phase of this project:

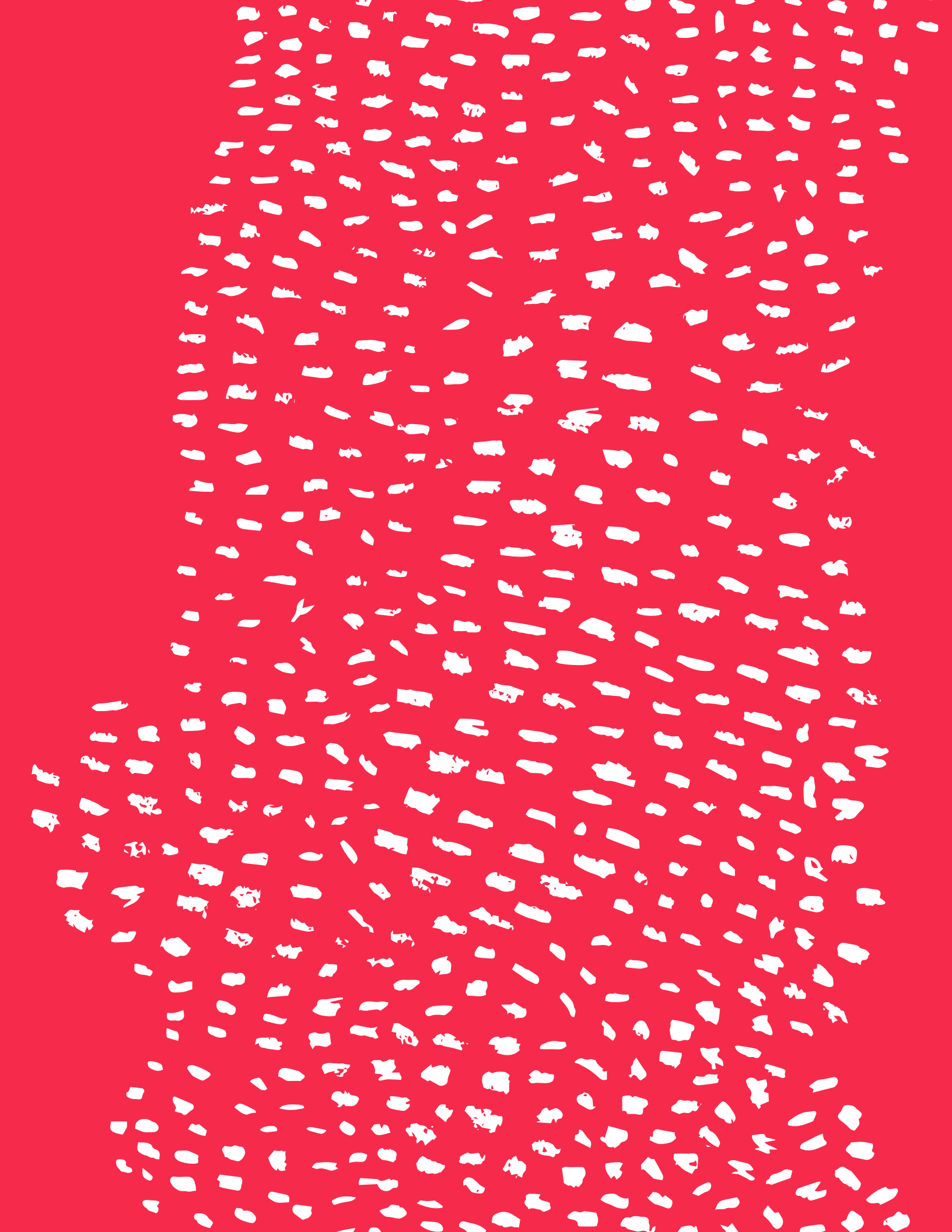
- Kristin Antin, HURIDOCS
- Jay Aronson, Center for Human Rights Science, Carnegie Mellon
- Patrick Ball, HRDAG
- Alexis Bautista, Migrant Forum in Asia
- Anh Bui, Benetech
- Neil Blazevic, East and Horn of Africa Human Rights Defenders Project
- Laura Carter, Amnesty Kristy Crabtree, International Rescue Committee
- Elsa Marie da Silva, SafeCity
- Priti Darooka, Programme on Women's Economic, Social and Cultural Rights
- Jessica Dheere, SMEX
- Nicola Diday, swisspeace
- Tarek Dobo, Syrian Justice and Accountability Center
- John Emerson, NYU Center for Human Rights and Global Justice
- Wael Eskandar
- Emmanuel Freudenthal
- Niamh Gibbons, Harvard Humanitarian Initiative
- Mahbul Haque, Bangladesh Centre for Human Rights and Development
- Morgan Hargrave, WITNESS
- Theresa Harris, AAAS
- Shevy Korzen, The Public Knowledge Workshop
- Tom Longley
- Milena Marin, Amnesty International
- Beatrice Martini, Aspiration
- Ruth Miller
- Tawanda Mugari, Digital Society of Zimbabwe
- Yvonne Ng, WITNESS

- Dan O'Clunaigh
- Ted Perlmutter, Institute for the Study of Human Rights at Columbia University
- Robin Pierro, European Inter-University Centre for Human Rights and Democratization
- Enrique Piraces, RightsLab
- Vanya Rakesh, CIS India
- Vijay Rao, Syria Justice and Accountability Center
- Anja Reiss
- Mike Romig
- Bridget Rutherford, PILPG
- Stephanie Seale, Benetech
- Marizen Santos, Migrant Forum in Asia
- Ryan Schlief, International Accountability Project
- Samaruddin Stewart
- Tom Trewinnard, Meedan
- Bert Verstappen, HURIDOCS
- Friedhelm Weinberg, HURIDOCS
- Eeva Moore
- Solana Larsen

Cover art: Lynne Stuart.
Graphic design by Federico Pinci.

First published June 2016

CONTENTS



Getting started

INTRODUCTION

From online videos of rights violations, to satellite images of environmental degradation, to eyewitness accounts disseminated on social media, we have access to more relevant data today than ever before. When used responsibly, this data can help human rights professionals in the courtroom, when working with governments and journalists, and in documenting historical record.

Acquiring, disseminating and storing digital data is also becoming increasingly affordable. As costs continue to decrease and new platforms are developed, opportunities for harnessing these data sources for human rights work increase.

But integrating data collection and management into the day to day work of human rights research and documentation can be challenging, even overwhelming, for individuals and organisations. This guide is designed to help you navigate and integrate new data forms into your human rights work.

It is the result of a collaboration between Amnesty International, Benetech, and The Engine Room that began in late 2015. We conducted a series of interviews, community consultations, and surveys to understand whether digital data was being integrated into human rights work. In the vast majority of cases, we found that it wasn't. Why?

Mainly, human rights researchers appeared to be overwhelmed by the possibilities. In the face of limited resources, not knowing how to get started or whether it would be worthwhile, most people we spoke to refrained from even attempting to strengthen their work with digital data.

To support everyone in the human rights field in navigating this complex environment, we convened a group of 16 researchers and technical experts in a castle outside Berlin, Germany in May 2016 to draft this guide over four days of intense reflection and writing.

THIS GUIDE IS FOR YOU

We assume you know how to do human rights research but wish to expand on your knowledge of how to use digital data and online media for documentation purposes.

This is a broad introduction that will set you on the right path to asking your own questions and seeking your own solutions. We aim to inspire critical thinking, rather than be prescriptive about what specific software, devices, or platforms should be used, since these evolve constantly.

We have vetted resources and collected further reading on the areas tackled in this report, which you can find online at ***https://engn.it/datnav***.

We imagine you are a human rights researcher, journalist, student, policymaker or philanthropist who wants to...

- Boost traditional research and documentation (interviews, surveys, and spreadsheets) by learning to incorporate digital data.
- Build knowledge and expertise in advance of the next emergency to avoid the reactionary collection of data while an event or violation is in full swing.
- Understand the opportunities, limits, and risks of digital data, as well as when and how to seek expert advice to help you achieve your goals.
- Overcome fear of digital data and technology, which is already in heavy use by your counterparts. With better tools, you know you can be more efficient.

NEW POSSIBILITIES WITH DIGITAL DATA

Some **forms of digital data** discussed in this guide:

- Photos, videos, and their metadata
- Satellite imagery and geospatial information
- Search engines, social media, and online opinions
- Government statistics and budgets

Today, millions of people have the ability to capture high resolution photographs or video with devices that fit into their pockets. With ease, they can share what they've captured with audiences near and far, projecting their observations across space and time, largely without regard to borders and language.

For human rights researchers, these new modes of information sharing are changing how we discover relevant information and reinforce the need for verification and healthy skepticism.

Compared to only a few years ago, there are immense possibilities that come from having so much data from so many sources to support human rights work. Digital data can support human rights documentation in depth, breadth, accuracy and efficiency.

First, it can offer a new source of corroborating evidence for traditional documentation of specific events. Human rights violations are often typified by an event. Armed with new data sources, the amount of detail that can be uncovered about a particular event grows (i.e. the who, what, where, and when can be identified more specifically), as does the impact of fact-finding.

Second—and perhaps even more significantly—data provides a standardised unit of measurement that can be captured, categorised and compared across groups and over time. This makes it helpful in mapping trends and patterns that help to spotlight more systemic dysfunctions.

Reboot of traditional methods

Human rights documentation has traditionally involved interviewing victims and witnesses and gathering corroborating evidence. It focuses on a specific event or incident. The goal is to find out what happened, to whom, by whom, where, when, and how.

Answering these six questions is fairly straightforward when governments violate a negative obligation (an obligation not to do something). Torture, arbitrary arrests, crackdowns on peaceful protests, forced evictions, and sterilization without consent are examples. However, the questions (what, to whom, by whom, where, when and how) are often much harder to answer for positive obligations.

Positive obligations require the government to do something. This could be to do something *at all* or to do *something differently*. A vast number of human rights violations—particularly violations of economic, social and cultural rights—fall into this second category. Human trafficking, labor exploitation, police brutality, malnutrition, homelessness, illiteracy, and preventable disease are examples.

These type of violations are complex and deeply-rooted. They're not limited to a specific event or incident. Instead, they result from systemic dysfunctions in ways laws, policies, and regulations are designed and implemented, which traditional documentation methods struggle to uncover. Those dysfunctions can be caused by a large number of actors and factors, which makes it hard to pinpoint who is responsible

A broader approach to documentation can help shed light on the actors and factors that influence the way laws, policies and regulations impact groups in particular ways.

Case study

Making better use of data from Syria to document violations

In April 2011, a group of activists in Syria began systematically monitoring and documenting the human rights violations by collecting and watching online videos of violence and atrocities. When several people from the team were kidnapped, the remaining members set upon improving their security, methods and datasets through consultation with experts.

They eventually developed a non-profit organisation called the ***The Violations Documentation Center (VDC)*** and now collect data on the imprisonment, torture, missing and killed civilians, rebels, and regime forces in Syria using rigorous verification methods, so their documentation could potentially be used in transitional justice events when the conflict ends.

They have more detailed database categories, more filtering options for search, and make better use of satellite imagery for corroboration. “Now our information is used confidently by UN representatives, governments and advocates around the world, since we feel certain that our data is complete,” says a VDC representative.

Photo: ***Damascus***, by ***Игорь*** M, CC-BY-SA 2.0 licensed.

Case study

Investigating mass killings in Burundi

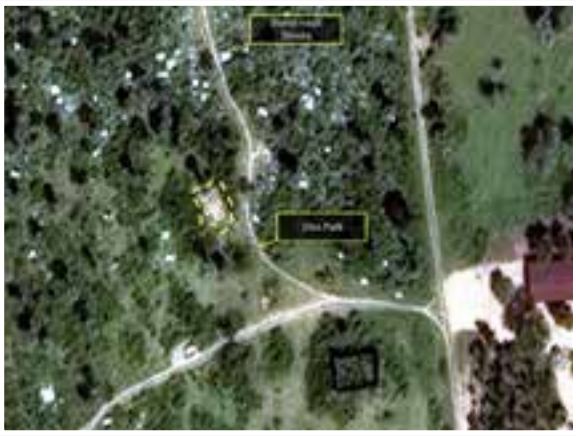
The bodies of at least 70 people executed by government forces in Burundi in December 2015 mysteriously disappeared leading to rumours of a mass grave.

An Amnesty International field researcher spent 10 days taking photos and interviewing witnesses after the killings, but more evidence was needed.

Satellite photos could probably help locate a grave, if only they knew where to look. Thankfully, Amnesty received a video from a contact in Burundi showing the alleged burial site. Google Earth was consulted online for the exact coordinates, and eventually satellite images really showed disturbed earth on the identified spot.

The finding of a mass grave via satellite gained massive media coverage and helped create political pressure. Ten years ago Amnesty would not have had such a video, since people in Burundi wouldn't have had cellphones with cameras. Thanks to the evolution of technology, investigators now have many more tools to incorporate in their traditional field work.

Photo: Satellite image showing disturbed earth in the Buringa area, which is consistent with witness accounts and video footage of mass graves. ©DigitalGlobe 2016.



For any data to be useful for human rights documentation, you need to be able to show the place it came from, when it was created, who created it, and why. Place, Time, Person, Motivation.

WHEN SHOULD YOU USE DIGITAL DATA?

If you are more accustomed to traditional research and documentation, it can be daunting to assess whether digital data or social media content is worthy of your attention.

This section lists seven questions to help you assess the net value of newer kinds of data in the context of your particular research. Will it really help? We need to constantly weigh the pros and cons, including the costs, savings, and ***risks for mishandling data***. For instance, the risk of surveillance can make digital data collection inadvisable in some settings, or you may only be able to access some populations but not others skewing the results of your investigation.

Really, it's not so different from traditional research methods. The hardest part is thinking creatively about what data may exist (especially data that was created for other purposes) and evaluating whether you and your team are able to collect and handle it effectively.

Seven things to consider before using digital data for human rights

1. Would digital data genuinely help answer your research questions? What are the pros and cons of the particular source or medium? What might you learn from past uses of similar technology?
2. What sources are likely to be collecting or capturing the kinds of information you need? What is the context in which it is being produced and used? Will the people or organisations on which your work is focused be receptive to these types of data?
3. How easily will new forms of data integrate into your existing workflow? Do you realistically have the time and money to collect, store, analyze and especially to verify this data? Can anyone on your team comfortably support the technology?
4. Who owns or controls the data you will be using? Companies, government, or adversaries? How difficult is it to get? Is it a fair or legal collection method? What is the internal stance on this? Do you have true informed consent from individuals?
5. How will digital divides and differences in local access to online platforms, computers or phones, affect representation of different populations? Would conclusions based on the data reinforce inequalities, stereotypes or blind spots?
6. Are organisational protocols for confidentiality and security in digital communication and data handling sufficiently robust to deal with risks to you, your partners and sources? Are security tools and processes updated frequently enough?
7. Do you have safeguards in place to prevent and deal with any secondary trauma from viewing digital content that you or your partners may experience at personal and organisational levels?

Practical examples

Here are some hypothetical examples to show a variety of digital data sources and uses.

Scenario

Refugees denied access to health care | Should we use digital data?

You wish to investigate allegations about denials of access to health care to refugees on the move and are considering using satellite photographs to map recent travel pathways and social media content to seek testimonials and interview subjects. You have a very limited budget and you haven't used this type of data before, but you know that most refugees have mobile phones and use social media to share information. Your biggest concern is that location data could land in the hands of local vigilante groups that attack migrants. In fact, because of vigilantes, refugees have increasingly moved their communications from more public platforms like Twitter to closed groups for chatting on Facebook and WhatsApp.

What do you do?

You decide to seek access to closed social media groups through your contacts and make security of everyone you communicate with online your highest priority. You query individuals and groups about denials of access to health care. For now, you forgo using aerial photos.

RESEARCH QUESTION:
Are migrants being denied access to health in Country X?

Government X is a point of entry to the European Union for refugees. There have been reports of access to health violation for pre-determination refugees entering by boat. This case study assumes that the researcher does not know exactly where refugees are entering the country.

HEALTH RIGHTS VIOLATIONS FOR REFUGEES

TYPES OF POTENTIAL DATA

- SATELLITE DATA
- GEOSPATIAL DATABASES
- WEATHER INFORMATION
- OPEN DATA[a]
- CROWD SOURCED INFORMATION[b]
- RANDOM SAMPLE OF DATA FROM MEDICAL CENTERS[c]
- 'CLOSED' SOCIAL MEDIA DATA[d]
- NON-SATELLITE AERIALS INFORMATION
- GOVERNMENT BUDGET INFORMATION[e]
- SOCIAL MEDIA DATA INFORMATION
- OCEAN DRIFT DATA[f]

a From humanitarian organisations and other potential partners (such as health care information, delivery analytics, etc).

b Such as citizen reporting maps of incidents related to refugees at entry point.

c Structures that provide medical support to refugees.

d For example Facebook or Whatsapp groups used by refugees to communicate re: route support or security concerns.

e Allocations to medical support of refugees.

f Or other factors that may a ect movement of refugees.

Scenario

Tracking Internet censorship | Should we use digital data?

You are working in a country where the government is known for shutting down the Internet during protests and social unrest and wish to document this. You discover that you can access ***Google's service disruption*** records online, which would allow you to identify likely down-times that can be cross-referenced with other additional sources of data. However, the government is adept at digital and Internet surveillance and could trace your browser activity and search history back to you, identifying you as a human rights activist.

What do you do?

Given the government's history of surveillance, your organisation has trained you to use ***VPNs (virtual private networks)*** for searching and communicating online. You decide that you can sufficiently hide your online tracks to do your research despite the risks.

RESEARCH QUESTION:
Did the government of Country Y violate freedom of expression and access to information rights by preventing access to the Internet?

There has been a wave of public protests in country Y for the last year. Recently this came to a head in a massive protest shutting down the downtown core of the capital city. Activists reported that the government blocked access to the Internet to suppress protests and prevent information from reaching out of country actors.

INTERNET RIGHTS VIOLATIONS

TYPES OF POTENTIAL DATA

- INTERNET SERVICE COMPANY RECORDS
- INTERNET TRAFFIC STATISTICS[d]
- MEDIA REPORTS
- SOCIAL MEDIA
- DOWNLOAD RECORDS OF NEW APPS[a]
- TELEPHONE RECORDS[e]
- INTERNET SPEED ANALYSIS
- GOVERNMENT BLOCKING ACCESS TO THE INTERNET[f]
- PRIVATE SECTOR/NGO STATEMENTS[b]
- PSIPHON DATA
- TAKEDOWN REQUESTS OR USER CLOSURE[c]

a In particular the ones that are considered safe and accessible.

b Statements from private companies (WhatsApp) & human rights organisations.

c Reported by social media or companies.

d Generally and on specific sites.

e Mobile phone usage data.

f In order to suppress protests.

Scenario

Violating access to water | Should we use digital data?

Aerial photos alongside rivers in a specific region show what appear to be physical barriers to accessing water. As the main source of water for the local population, you are concerned that the company is violating the water access rights of locals. The population living along the rivers tends to be poorer, though most have mobile phones. You would like to send out an SMS survey to ask if and where their access is being limited, but the population has been the target of a lot of development work in recent years and, you have heard, is suffering somewhat from survey fatigue. Your NGO is young and has limited resources.

What do you do?

Even though the SMS survey will enable you to reach your target population, the context may prevent you from getting responses. Given that (1) you are targeting a poorer population with survey fatigue, (2) you will have to ask each respondent to pay for the texts herself or himself, and (3) your NGO has very little name recognition, you realise this might be an uphill battle. You decide to ask a colleague at a well-established NGO to introduce you to a community leader so that you can explain your project, ask for the community's input in designing the survey, and figure out how to feed the information on your evolving project back to the community.

RESEARCH QUESTION:
Is FPower responsible for the pollution of the water supply?

In country X there are high levels of malnutrition amongst low income children due to dysentery. There are also a lot of commercial ower farming operations located close to rivers throughout the country. The majority of the farms are owned by one conglomerate company; FPower. The Board of FPower has two members who are directly related to government ministers.

ACCESS-TO-WATER VIOLATIONS

TYPES OF POTENTIAL DATA

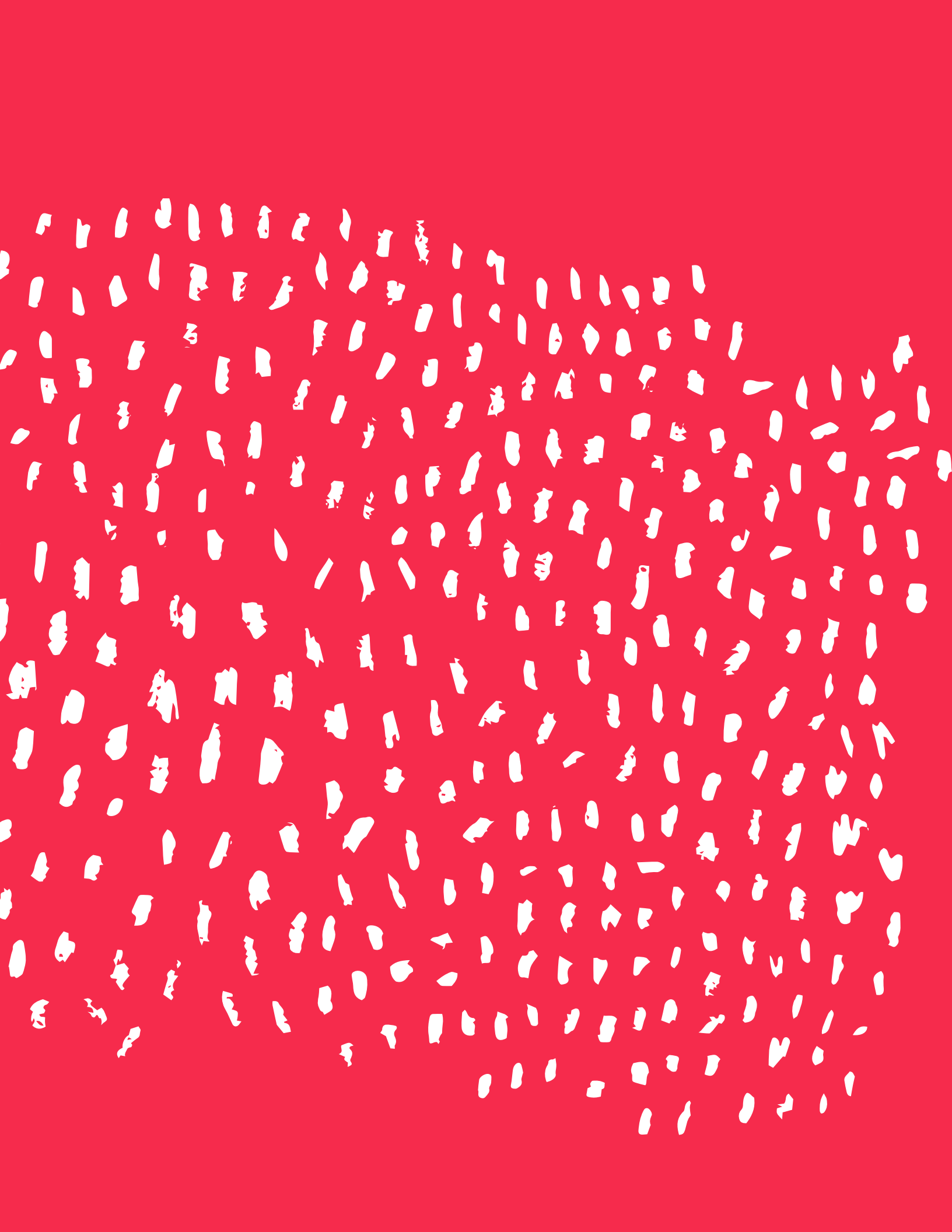
- NATIONAL HEALTH STATISTICS
- PHOTOGRAPHS OR SATELLITE IMAGES[a]
- INTERNET SEARCH/ DATA/ ANALYTICS
- BUDGETARY DATA[b]
- BUSINESS LOCATION AND OPERATIONS INFORMATION
- SERVICE DELIVERY DATA[c]
- SMS SURVEY ABOUT CALORIE CONSUMPTION
- WATER POLLUTION ANALYSIS
- OVERLAP OF GEO AND TEMPORAL DATA

a Time lapse or geospatial.
b Government spending on nutrition programs.
c Government or humanitarian agencies.

OBTAINING ORGANISATIONAL SUPPORT

Even where there is appetite for change, long hours and tight budgets can keep organisations and individuals wedded to the old model of doing things. In order to incorporate modern data practices, therefore, organisational leadership needs to champion change. Here are **ten points to include in a discussion with managers**, to help them see the necessity of change.

1. **Market Reality**: Many human rights organisations are already working with digital data, but many are doing so without investing in data-literacy. Failure to develop the necessary in-house skills puts the relevance of your work at risk.
2. **Budget**: Many of the tools recommended in this guide are low-cost or free.
3. **Training**: Yes, training takes time, but employees value training. Research shows that employees value new training opportunities at work second only to compensation. Especially for small organisations, keeping staff and reducing turnover is vital. Places like the School of Data or Advocacy Assembly offer free online courses.
4. **Attractiveness to new hires**: Being a forward-looking organisation that utilises new tools helps attract the best people in the field.
5. **Impact**: Using new data effectively can improve outcomes and enable better monitoring of human rights violations.
6. **Funding**: Donors are keen to link new calls for funding proposals to digital data and their use in human rights reporting.
7. **Branding**: It makes a brand look smart, forward-looking, and willing to engage with new developments.
8. **Mission**: The mission is to monitor, report, and raise awareness of human rights violations. Used correctly, digital data allows organisations to do this better.
9. **Lunchtime talks for staff**: Invite someone who works successfully with digital data to give a talk about what they are doing.
10. **Build peer networks**: Build relationships with data-oriented peer organisations to share best practices.



Understanding verification and documentation

WHAT IS METADATA?

Metadata is information about a file (such as a word document, a PDF, a picture, music file etc.) that is stored within the file itself, hidden from view.

This information can include the time and date a file was created, the username of the people who created or edited it, information about the device that created it, and other kinds of information. In other words, the metadata in a file could expose who created a file. This information is generated automatically by devices like cameras, computers, and phones, but it can also be edited and manipulated by those who know how. This can be a good thing, if you want information to be private when you share files, but the risks are obvious if you are dealing with verification of sensitive human rights information.

Let's take pictures as an example. When you take a picture with your digital camera, what happens? If your camera or phone knows where you are, then that information (in the form of GPS coordinates) can be saved in the metadata of the file. If your camera knows what time it is, it records the date and time the picture was taken. If your camera or phone has a serial number that may be recorded in metadata as well.

Digital image file formats such as TIFF (Tagged Image File Format) and JPEG (Joint Photographic Experts Group) created by digital cameras or smartphones, contain metadata in a format called EXIF (Exchangeable Image File Format) which could include all of the above information and even a thumbnail of the original image.

Other files such as text documents include metadata too. It could be about how long the document is, who the author is, when the document was edited, and a short automated summary of the content. EXIF data and other kinds of metadata can be extremely useful for verification of media, and also for the creation and organisation of catalogues.

Scrubbing metadata

In some cases, you may want to remove metadata from files related to human rights violations. This is particularly relevant in cases where sharing or publicising evidence of human rights violations carries a threat to you or others involved in recording the incident.

A number of software and online tools can be used to 'scrub' metadata from files. While not all metadata can be removed – such as the size of a file, the dimensions of an image, or the time it was last edited–metadata related to who created and edited the file can often be removed. The efficiency of tools varies, so it is best to try a couple and check that the metadata is effectively removed.

If you are working with a trusted correspondent, you may suggest that they turn off location services (GPS data) on their device to hide their identity. Exercise caution when requesting images, since the phone model or other EXIF info could be used to locate your sources.

The guide ***Security in a Box*** from Tactical Technology Collective and Frontline Defenders has even more information on removing metadata.

VERIFY, VERIFY, VERIFY

Data of any kind–even materials you've collected yourself–require thorough verification to protect your reputation and people on the ground from suffering harm. It is important to approach each piece of content with a critical eye, even if you wish for it to be true.

You also have to accept that there is no single method to guarantee verification. Verification is a process for giving yourself and others confidence in the veracity of the content. You should be as transparent about what you don't know as what you do know, for everyone's sake.

In high-pressure, low-resource environments verification of digital data is often an afterthought rather than something that is incorporated in research plans from the start. If you have doubts about how to verify a particular type of content, it's safer to consult experts to help you.

Five steps of verification

1. **How did you get the content?**
   Think about what information channels it traveled through before arriving on your desk. How many times did it change hands?
2. **Who created the content?**
   Is the person who shared or uploaded the content online also the creator, or was it someone else? Ask if you don't know.
3. **Where is the content from?**
   Descriptions and metadata can easily be forged. Are their visible landmarks or sounds (like police sirens or dialects) that can help you verify a location?
4. **When was the content created?**
   You can't trust the date stamp on a file. Are there visual clues like the weather? A reverse image search can show if a photo appears elsewhere.
5. **Why was the content created?**
   Can you determine the motivation for sharing the content? What interests does the uploader have?

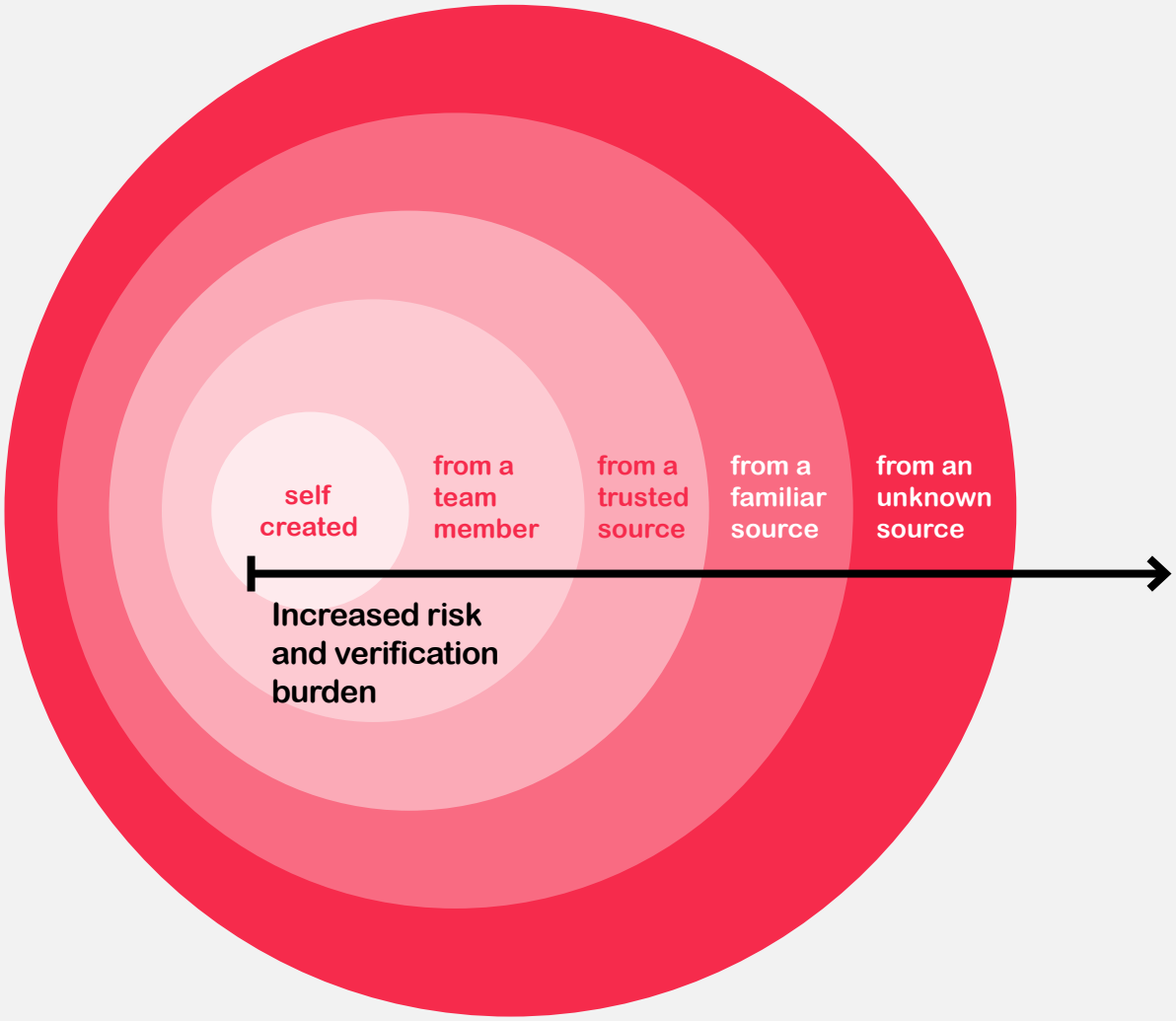
Why Verification?

It may sound obvious, but unless working with materials you've collected yourself or your own firsthand observations, it is important to verify that the information you are working with is, in fact, what it purports to be. Reputational risks, as well as risks to the subjects of the research, require due diligence when working with new data of any kind.

Verification steps

Verification is an evolving process. And the end product is rarely definitive. Rather it is a process that lends confidence in the information. This information may purport to tell you something about the who, what, where, and when of a certain event. Asking the same questions of the information itself is part of the verifications process. Thinking about these questions will define which tools you need to engage to verify the piece of content:

WHAT is the source?
WHO uploaded or shared the content?
WHERE was the content created?
WHEN was the content created?
WHY was the content created?



Exercise

Verify this YouTube video

You are a human rights researcher, and need to verify a YouTube video [*https://youtu.be/clrICxJihWI*] of excessive use of force by police at a protest in Rio de Janeiro, Brazil that a colleague shared with you.

1. You **download the video** to *preserve it*, because sensitive online content gets often removed.
2. Using the *YouTube Data Viewer*, you **identify the time of publication** (though not the time of the recording) which was 5:30pm (Brasilia Time) time on April 14, 2014. A reverse image search on thumbnails of the video, does not retrieve any earlier versions of the video online.
3. The creator helpfully listed their name under the video, and you conclude by looking at other social media platforms that **the uploader** appears to be an activist in Rio de Janeiro.
4. The **video and description** is rich in clues about the location (Prefeitura do Rio). To verify, you use the online map *Wikimapia* to find Rio's city hall (prefeitura) and use satellite images found on Google Earth to cross-reference visible landmarks (office building, pedestrian bridge). Geo-referenced photos from *Panoramio* (accessible through Google Earth) also match up.
5. The description says the protest happened on Monday morning in heavy rain. *Historical weather data* accessed through *WolframAlpha.com* is consistent with what is seen in the video (rainy and misty) thus making the video and **the alleged time of day** seem more credible.
6. You use slow-motion video playback (on YouTube or *VLC media player)* to take a closer look at police **uniforms and insignias** compared with those depicted on *official police websites* [*https://www.facebook.com/gmrio.oficial*]. You note down details, in case it can be useful for future investigations of individuals involved.
7. Finally, when you search for additional content with your regular search engine, you are able to find several **other videos from the same date and protest** [*https://youtu.be/sxyrP0yBBts*], corroborating the video.

How to verify images using EXIF files (Q&As)

Each and every digital photo contains a metadata file called EXIF that can identify when and where the image was captured. This information is added to the file the moment you take the picture, no matter what device you use. EXIF files can be very useful for verification, but can also be manipulated when an image changes hands or has circulated on social media.

When should I look for an EXIF file?
A picture should contain an EXIF file if it came directly from the source's camera to your inbox without being altered or changing hands. It can still have an EXIF file if it was edited. If an image does not contain an EXIF file, you should question the originality. You can check the EXIF and other data and see what kind of camera created it or what software it was exported from. If it has been altered or manipulated, it may have other kinds of metadata as well.

Will there always be EXIF data in your image? Most social media platforms strip out EXIF data or create lower quality copies of the photo when a picture is uploaded to their platform (with the exception of dedicated photo sharing websites). If you're verifying a picture sourced from social media, you usually won't find any EXIF data.

How do I find EXIF data?
If you search online for "EXIF data viewer" you will see many options. One of the simplest tools is Jeffry's EXIF Data Viewer [*http://regex.info/exif.cgi*] which also offers a plug-in for several web browsers. You can also find EXIF data using software on your computer like Photoshop or iPhoto.

Can EXIF Data be manipulated?
Yes. Tools such as ***Geosetter*** and other photo-editing software can be used to falsify EXIF data. This means that EXIF data should only be used as one of many steps in your verification process. Tools like ***JPEGSnoop*** may detect software that was used to manipulate an image.

What if someone takes a photo of an old photo?
EXIF data only tells you the device that captured the image, it knows nothing about what was actually in front of the camera. Even if the EXIF file matches up to date and location, the image could still be something other than what it pretends to be, for instance a photo of a photo.

In depth

What a photo says about you... This is what EXIF files look like

Basic Image Information

Target file: 154ec836b067df8d25e1.jpeg

Camera:	Htc One X9 dual sim
Lens:	3.8 mm
Exposure:	Auto exposure, Not Defined, 1/10 sec, f/2, ISO 800
Flash:	none
Date:	**May 26, 2016** 12:01:45PM (timezone not specified) (8 hours, 56 minutes, 50 seconds ago, assuming image timezone of US Pacific)
Location:	Latitude/longitude: **53°** 13' 54.9" North, **11°** 51' 1.2" East (53.231922, 11.850347) Location guessed from coordinates: *K7044, 19348 Berge, Germany* Map via embedded coordinates at: Google, Yahoo, WikiMapia, OpenStreetMap, Bing (also see the Google Maps pane below) Altitude: 0 meters (0 feet)
File:	**2,368 × 4,160** JPEG (**9.9** megapixels) 2,316,219 bytes (2.2 megabytes)

Only some EXIF data will be relevant for verification.

1 **The camera used to capture the image**

This is the model of camera or phone. If, for instance, you are corresponding with a person who has shared an image with you, you can corroborate what the EXIF data tells you about the camera used to capture the image with what the person tells you about the device they have captured it with. If you're told a different camera or smartphone than shown in the EXIF data, that's a big red flag.

2 The date and time the image was captured

This is important for checking the time a picture was captured, but EXIF does not indicate time zone. It can either reflect local time or Coordinated Universal Time (UTC), depending on device and settings. Some camera brands have their own metadata tags that supplement EXIF, and may include time zone data. There is also a non-standard EXIF tag for TimeZoneOffset, which can indicate the time offset relative to GMT. Many media upload sites, such as YouTube, default to Pacific Coast Time if the timezone is not embedded in the image. You should also consider if the time and date settings of the device could have been incorrect. This is more likely with a camera than a smartphone, which often auto-updates time and date settings.

3 GPS co-ordinates of where the image was captured

If the device capturing the image has GPS capabilities (mostly smartphones, but some cameras do too), and the person capturing the image has switched on these capabilities, the EXIF data may show the GPS coordinates of where the image was taken. GPS stands for Global Positioning System–a satellite location system. If available, most EXIF data viewers will show these coordinates on an online map service.

4 The pixel dimensions of the image

The pixel dimensions (width and height) is helpful in two ways. First, different cameras that produce images of the same pixel count can produce images with different dimensions. For instance, Fujifilm XT-2 and Leica M-D Typ 262 cameras produce 24 megapixel images, but with different pixel dimensions. The Fujifilm XT-2's pixel dimensions are: 6000 x 4000 pixels. The Leica M-D Typ 262's pixel dimensions are 5976 x 3992 pixels.

Pixel dimensions may also provide clues if an image has been cropped. As a rule of thumb, the width and height of an image should result in a whole number when divided by eight. The resolutions of the majority of digital cameras are multiples of eight, which can be verified by reviewing a camera's technical specifications. (There are exceptions to this rule, especially with mobile apps and panoramic functions).

5 The pixel count of the image

Each camera shoots at different image sizes. For instance, an iPhone 5 produces an 8 megapixel image, while a Samsung Galaxy S5 produces a 16 megapixel image. In the graphic here, we see that the image is 9.9 megapixels , i.e. 2368 x 4160 pixels (one megapixel = one million pixels). You should raise a red flag and ask questions of your contributor if the pixel count and dimensions do not match the sizes supported by the purported device.

Using reverse image searches to verify and identify

Image search engines are great tools for verification and identification. When you upload a photo to an image search engine and perform a "reverse image search" it checks for matching file names or similar pictures that already appear online. This can help you determine whether a photo is what you think it is, older than you believe, or was used previously in a different context or country.

1. **Find the right search engine for you.** *TinEye* and *Google Image Search* offer the most comprehensive databases. But many other search engines (Bing, Yandex, Baidu) also offer reverse image search. Test to see which one works best in a geographic context.
2. **Always perform two searches.** Each reverse image search engine queries a different database and indexes new images at different speeds. Your image may appear in one search engine while not in another.
3. **Order your search results by oldest first.** If the image is from a different event than is claimed, this will quickly reveal the discrepancy.
4. **Locate landmarks geographically.** If you are trying to identify a landmark in an image or video, a search engine that enables "similar to" searches is a great help, as landmarks are highly documented.

Beware! A reverse image search cannot tell you when an image was taken. It only tells you if an image has been previously indexed on the Web and when.

Beware! If your image does not appear in a reverse image search, it does not mean it is new. The image could have been sitting on a hard drive for years.

Case study

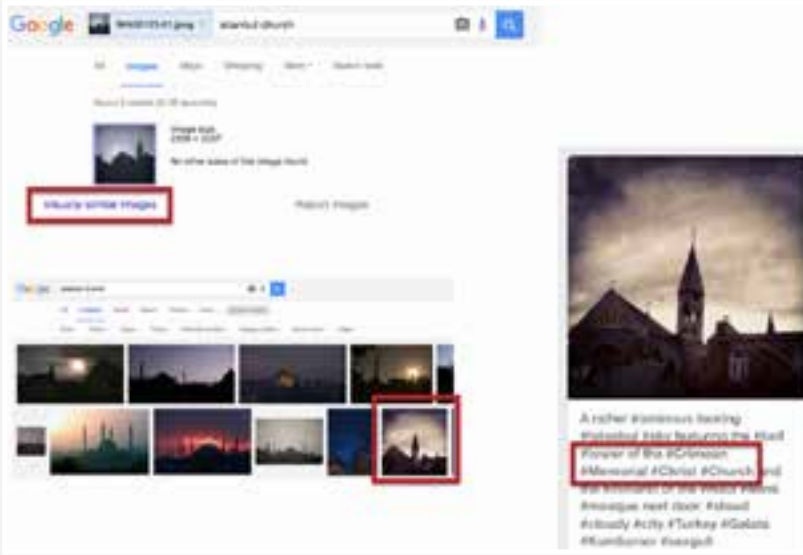
Verify the location of this image

We have a photo of a church which is said to be in Istanbul, Turkey.

To verify this landmark through reverse image search we upload the image to Google Images and type “Istanbul church” alongside the image.

We have the option of searching for “visually similar images”, which takes us to a page where one image has the same tower and roof. Clicking on that image suggests it’s the Crimean Memorial Christ Church in Galata, Turkey, which we can now check on a map.

You can also do this kind of verification with videos. Simply take a screenshot of a landmark in a video and follow the same steps.



SOCIAL MEDIA DATA

In June 2014, a video of an extrajudicial execution in the Central African Republic was posted online by a newspaper and was widely shared through social media. However, after verifying it, Amnesty International discovered that the incident actually took place in Nigeria, and the video ended up being the basis for Amnesty's research ***implicating Nigeria's military in war crimes***.

This example shows the enormous pitfalls and opportunities of social media content for human rights research. Whether text, photos, video or audio, it has the potential to uncover hidden atrocities, while at the same time endangering the reputation of an organisation if used carelessly. Additionally, the large amount of content shared across multiple platforms can easily overwhelm researchers and sometimes calls for specialised organising and discovery skills.

Social media is very different from traditional sources such as eyewitness testimony or newspapers, not least because the original source of information is often unclear, and because ***misinformation can spread*** incredibly fast from one person to thousands. At the same time, social media is growing increasingly important. Even governments and militaries use online channels (sometimes exclusively) to distribute information directly to the public. Recent examples include the Russian military's ongoing reporting on ***military operations*** in Syria on YouTube, and the ***Israeli Defense Forces*** online broadcasts about operations in Gaza in 2014.

A couple of examples can help illustrate the vast opportunities for event reconstruction and long term investigations. In ***Egypt***, content from social media allowed reconstruction of the killing of a peaceful protester by police in 2015, and subsequently forced an investigation and trial. In ***Nigeria***, social media content played a crucial role to document war crimes in 2014.

When researchers are more skilled in asking questions of social media and employ technical tools to parse bigger data sets, they can do advanced analysis across time, place or language. In 2016, ***Abodo***, an online housing marketplace, ***analysed millions of tweets*** to see how use of derogatory language about race, ethnicity, gender, religion, and sexual orientation varied in different parts of the United States. In 2014, Global Voices performed ***a "sentiment analysis" of Russian tweets*** about the threat of war in Crimea showing weak opposition.

Getting started with social media research

Always use social media for human rights advocacy in a secure, ethical and impactful way. You need to *discover*, *organise*, *preserve* and *verify*. The following questions are intended to provide guidance. Use them in your research, regardless of what platforms you are working with.

What platforms and apps are dominant in your country or region?

"Social media" is a broad term, and user applications vary across countries and regions. While Twitter may be popular in the United States, people in Egypt or Burundi may prefer to use Facebook. Recognising these differences is crucial when tracking human rights incidents.

What keywords and hashtags are used to share content on your topic?

People communicate using specific keywords or hashtags. Follow these identifiers to discover content. Eg. users in Burundi used #1212massacre to describe violence of December 2015. Also, language used by witnesses may be more emotive than what you first search for.

Can you narrow your search and filter results?

Some social media platforms allow search by location, date or type of content. For instance, ***Tweetdeck***, a tool for managing Twitter, enables you to search for images and video only, while excluding retweets. You can also create ***lists*** of trusted users to monitor content selectively.

Photo: examples of the kind of words used on social media in breaking news situations, including "I" and "my" that could help you search for eye witnesses. **Source**: Deutsche Welle/Reveal Project

Have you saved the content you are reviewing?
Online content can disappear very quickly. It can be removed by the uploader or by the social network hosting the content. Take snapshots of postings or download images.For videos, there are tools like *VideoVault*.

Can you verify the information?
Social media messages have to be carefully vetted for accuracy. The temptation to use widely shared visual content uncritically could be higher than with eyewitness testimony, but the same verification criteria should apply.

Do you have other information sources for corroboration?
Do not rely on a single social media message when it is likely that other online postings, videos and photos reference the same incident. Combining different pieces of information, including eyewitness testimony and news reports, will help you to build a much stronger case.

Are there privacy risks or ethical concerns?
As with any source, consider the potential risks to individuals if you share or publicise their words or images. Also, ask yourself if you may be doing unintended harm to communities affected by a conflict.

In depth

Do-No-Harm

The Do-No-Harm (DNH) principle is a useful framework to keep in mind in the context of the peaceful transformation of conflicts (armed or otherwise). DNH requires posing questions about whether interventions contribute to **unification** or further **division**.

Sharing powerful images or video footage of human rights abuses or atrocities via social media could deepen divisions between two sides in a conflict, depending on how it is presented. If falsified material claiming to show atrocities spreads, the conflict could even intensify. This raises serious ethical concerns about sharing social media content and highlights the paramount importance of verification and thoughtful presentation of data, with a view to ending rather than perpetuating violence.

For more on the Do-No-Harm principle see *Anderson, Mary B: "Do No Harm: How Aid Can Support Peace – Or War" Boulder, CO: Lynne Rienner Publishers, 1999*.

Photo: "*Non-Violence*", UN Photo/Michos Tzovaras

Three categories of "fake" to watch out for in social media

Misattribution

This is by far the most dominant challenge in human rights research. Content is recycled online on an ongoing basis, and is shared with a wrong date, location or attribution.

Example

A highly graphic video was shared during post-election violence in the Ivory Coast in 2011. However, the video was ***several years old***, and had already been shared in multiple countries. A simple technique to detect old content is to do a ***reverse-image-search***. This can also be done with the preview image of a video. (See verification chapter of this guide).

Staged

A scene or event, or specific details of an incident, can be staged. For example, an armed group in Syria posed in a YouTube video, with ***what turned out to be toy guns***.

Example

The "Syria Hero Boy Video" was a ***staged video***, produced by a Norwegian filmmaker. It is a textbook example of how looking at the source and digital footprint of the original uploader should raise questions. The YouTube channel that ***first hosted the video*** was new and contained only this one, highly dramatic video–a red flag.

Technical manipulation

Be mindful of the increased ease with which content, especially images, can be manipulated by cropping, erasing details or merging photos to misrepresent events.

Example

In 2013, officials from the province of Anhui, China released a doctored photo which showed a vice-mayor "***hovering like a ghost over a puppet-sized elderly lady***." After a lot of online mockery, they admitted it had been the result of two merged photos and expressed "deep regret".

DATA THAT WILL HOLD UP IN COURT

Legally admissible evidence

Human rights documentation has a powerful impact on legal proceedings. The International Criminal Court, as well as several hybrid tribunals, actively seek civil society's support. But courts are slow to adopt new technologies, and judges, who tend to be older, are often not familiar with how digital data can be folded into normal evidentiary rules. It will vary from jurisdiction to jurisdiction whether a YouTube video will be accepted as legal evidence.

For example, in 2015, Sweden's justice system ***prosecuted a former fighter from Syria*** for torture based on a video posted to Facebook. Many Syrians are crowdsourcing such online content to prosecute abroad former Syrian officials, militaries, and fighters that have fled the country.

Requirements for court

If you want your data to lead to criminal prosecutions, you need to make sure to record all relevant metadata. A court should be able to clearly identify the location of a video, either through location coordinates embedded in the video or landmarks that can be identified in the video content. Moreover, the subject of the video should be clear – blurry or shaky video that casts doubt on who is in the film or what act they are committing will likely not be accepted.

Also consider chain of custody. To be accepted for a criminal prosecution, most courts need to know every individual who handled the content, starting from the original creator and ending with the prosecutor in the courtroom. Chain of custody must show exactly who possessed the content, when, and for how long in order to help safeguard against manipulation of evidence. The closer you get to the original video, the easier it will be to present in court because the chain of custody will be shorter and easier to demonstrate.

It will also be important that you work impartially and conduct very minimal tampering of the data. A defense lawyer might argue that the data is corrupted, biased, or carries improper motivations. You will need to conduct at least some analysis in order to make your documentation searchable and understandable, but try to keep the analysis at a minimum.

Risks of engaging with courts

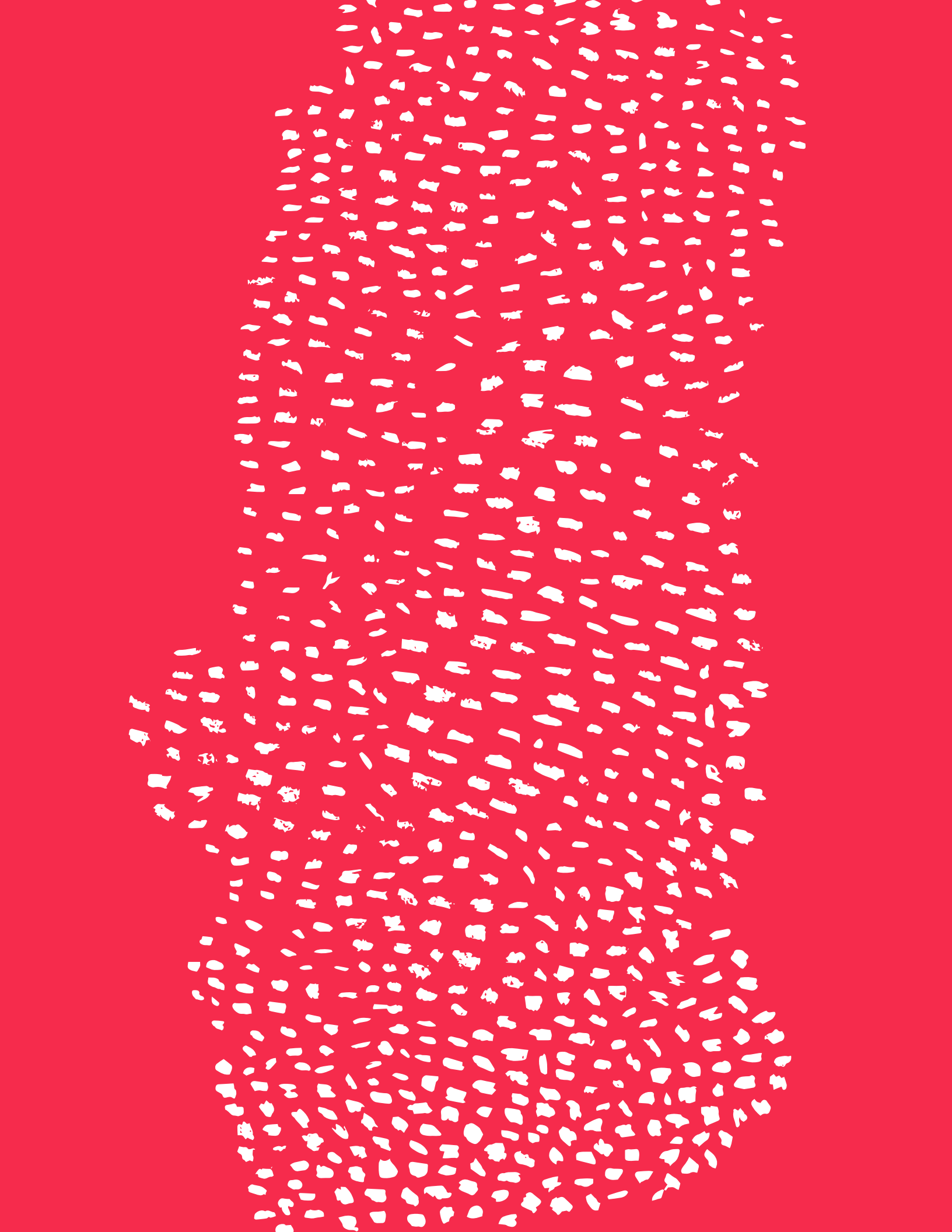
By engaging with courts, human rights organisations risk being subpoenaed to divulge confidential evidence, like names and dates, in open court due to a pre-trial procedure that enables each party to see evidence of the other ("rules of discovery").

If you are working with the prosecution, try setting up a witness protection system. Many court proceedings are public, and all your precautions could instantly go out the window if the witness testifies and their name and personal information go public. Avoid engaging with courts that do not protect witnesses so your witnesses are not harmed as a result of their testimony.

Other forms of justice

If these risks are overwhelming, all is not lost. ***The Syria Justice and Accountability Centre***, an organisation collecting human rights and international humanitarian law violations in the Syrian conflict, has collected over one million pieces of data over five years. Most of the documentation — perhaps up to 90 percent — will likely not be accepted in court, but all the remaining documentation can be used for other types of transitional justice processes.

Transitional justice includes truth commissions, reparations programs, memorialization, and institutional reform, which can all be used to help heal, redress past harms, and move towards reconciliation. In the case of institutional reform, large data sets can show abusive and corrupt trends in a country's judicial sector. This data can help demonstrate aspects of the sector that require reform and can guide a country's legal reform and vetting process. By reforming the justice sector, a country can move towards ensuring non-recurrence.



Practical techniques

OPEN GOVERNMENT STATISTICS

Governments collect data to inform decision-making and planning. There is a push for governments to open their data to the public to increase transparency, accountability and citizen engagement. Examples of open data initiatives are on the rise worldwide, including online portals and apps for citizens to access relevant government data for their own use. These sources can provide a wealth of relevant information for human rights researchers, but also poses challenges.

Accessing government data

A large amount of secondary data is publicly available on national bureaus of statistics websites. Additionally, many agencies publish government-programme evaluations online.

Where data is not available online, you may be able to submit a freedom of information request. More than 95 countries recognise a general right to access government information, with exceptions for information deemed too sensitive to share. Read more at: ***http://right2info.org/***.

Downloading, storing and organising government data

Government data is often presented in PDF format, making data searching and analysis challenging. The School of Data has tutorials on how to extract data from a PDF, so that you can work in a user-friendly and familiar format like CSV or XLSX.

Types of government data

Data type	Examples	Pros	Cons
Administrative records Created when government agencies and institutions interact with the public. They can also include transactional data on service providers and finances.	Vital statistics on populations, like death and birth rates; employment and job seeker statistics. Beneficiaries of social policies and services. Vendor contracts and bookkeeping.	Where administrative records are frequently updated, it can be a simple, trackable, quantitative resource that can improve transparency and help root out corruption.	Only recorded for people who use a particular public services, so coverage is not always reliable. For example, crime statistics may underrepresent sexual assault, which is extremely underreported.
Statistical surveys Also called a sample survey, statistical surveys only collect data from a subset of the population, with the goal of drawing inferences for the entire population.	Demographic health surveys, labor force surveys, household income and expenditure surveys.	Surveys can be a cost-effective way for governments to collect information where data from administrative sources are not available.	Selection bias may lead to the sample surveyed not being representative of overall population.
Censuses Catalogue of all members of a country or territory.	Countries usually conduct censuses of population, housing, agriculture and industrial establishments.	Censuses provide baseline data on key characteristics of the population and on variables that do not change rapidly.	A population census is usually conducted at ten-year intervals because of the complexity and cost of the operation.

Assessing the reliability of government data

Government data is not always accurate. It is not uncommon for governmental and nongovernmental databases to contain conflicting information. Differing methods of data collection, analysis and computation are leading causes of such conflicts.

Concept	Definition	Potential issues
Validity	Data must reflect what it is trying to measure (e.g. the fulfilment of a right) as closely and as accurately as possible.	Most secondary data is collected for uses other than human rights monitoring and so must be repurposed. This reinterpretation can be challenging.
Reliability	Refers to the consistency or dependability of the data. In other words, data collected multiple times in the same way should deliver similar results.	Ambiguities or biases in the way data is collected (e.g. how survey questions are framed or how a population is sampled) might make data unreliable.
Impartiality	Data must be collected in a way that respects scientific independence and in an objective, professional and transparent manner.	A national statistical office that is not independent may "spin" numbers so that a situation appears better than it is.

To assess the validity, reliability and impartiality of the data, ask:

- **What is the subject of the data?**
  On some particularly sensitive or controversial topics, government data can be notoriously unreliable.
- **How are survey questions framed?**
  Sometimes questions are leading, and therefore bias the answer.
- **What is the sample size?**
  It may be too narrow to be representative.
- **How frequently is the data collected?**
  The data may be stale.
- **Who collects the data?**
  Potential conflict of interest or resource limitation.
- **Who publishes the data?**
  Potential conflict of interest.

Analysing statistics from a human rights perspective

When combined with and analysed against human rights standards and principles, government data can be useful in exposing violations of economic, social and cultural rights. When looking at government statistics, therefore, it's helpful to ask: What does this tell me about the availability, accessibility and quality of goods and services? Are there particular regions or particular groups that are marginalised or discriminated against? How have things changed over time? Have they gotten better or worse?

Case study

Government data as evidence of systematic discrimination in Guatemala

In 2009, the Center for Economic and Social Rights (CESR) and the Instituto Centroamericano de Estudios Fiscales published a report on the right to maternal health in Guatemala that showed proof of discrimination against indigenous women.

Relying on data from the World Bank, United Nations Development Programme and national statistics bureaus, they cited demographic health surveys showing that Guatemala had among the worst and most unequal maternal mortality outcomes in Latin America. Indigenous women were three times more likely to die during childbirth or pregnancy than non-indigenous women, and over 50% of deaths could have been preventable with skilled care. Administrative statistics showed serious issues with the availability and quality of services.

The poor implementation of policies related to inadequate resource investment in the health sector. Allocations to health had remained around 1% of GDP since the end of the war in 1996, lower than those of poorer Central American countries. Distribution of per capita health spending was also highly inequitable, with three times more money going to the capital than to Quiché, the poorest region. Low social spending was directly linked to the country's low tax base, which was mostly generated through regressive indirect taxes that disproportionately target the poor while the country's business sectors enjoyed tax privileges and incentives.

All of these figures combined to build the strong conclusion that Guatemala was not doing all it reasonably could to improve maternal health, in effect discriminating against poorer, indigenous women. See full report here: ***http://www.cesr.org/section.php?id=33***.

Photo: ***Street Scene, Chajul, Quiché*** in Guatemala (2014) by Adam Jones on flickr (CC-BY-SA 2.0)

BUDGET DATA FOR HUMAN RIGHTS

In setting budget priorities, governments can unwittingly or deliberately disregard human rights standards. For example, budget cuts to the criminal justice system may leave low-income defendants in pretrial detention for unreasonably long periods. A decision to reduce subsidies or increase taxes on particular households or hygiene products may indirectly discriminate against women. Budget data literacy can, therefore, be an essential component of comprehensive human rights research, especially now that so much data is available online.

Types of budgetary data

It can be helpful to break a budget into three main parts: how revenue is generated, how budgets are allocated and how expenditures are actually made.

Data on budget revenue

Are tax policies generating *sufficient* resources? Are they doing so *equitably*?

Helpful data for judging the *sufficiency* of resources includes:

- Government revenue as a percentage of GDP.
- Tax revenue as a percentage of government revenue.
- Tax effort (the ratio between the actual tax collection and taxable capacity).
- Volume of illicit financial flows.
- Tax revenue as a percentage of total tax revenue.

Helpful data for judging the equity of resources includes:

- Composition of taxes (e.g. percent made up of income tax, taxes on goods and services, corporate tax, etc.)
- Tax as a percent of overall income paid by different groups.
- Exemptions granted.

Case study

Money down the drain

A country allocates 1.5% of its budget to the sanitation sector. This allocation has decreased over the past ten years. 75% of the money it allocates to sanitation subsidizes waterborne sanitation (i.e. sewerage pipes), but poor households in informal settlements rely on on-site sanitation (e.g. pit latrines). Does this raise red flags from a human rights perspective? The government may, in fact, be discriminating against households in informal settlements.

The International Budget Partnership has a series of resources that explain how international human rights standards like progressive realisation, non-discrimination, and maximum available resources can help answer this question: ***http://www.internationalbudget.org/publications/escrarticle2/***.

In depth

What is gender budgeting?

Gender budgeting is a particular type of budget analysis used to assess the impact of government revenue and expenditure on women, men, girls and boys. For example, in the area of health, men and women have similar needs in respect to influenza and malaria. However, women have greater needs than men in terms of reproductive health. Gender budgeting is a new and evolving way to visualise and address the discriminatory effects of resource decisions. There are helpful resources available at: ***www.gender-budgets.org***.

Photo: ***San Malen Primary Health Unit in Pujehn, Bo district, Sierra Leone*** (2013) by H6 Partners on flickr (CC-BY-NC-ND 2.0)

Data on budget allocation

The following steps can determine if budget allocations are in line with human rights standards and principles.

1. **Calculating**
   - *Ratios* or *shares* (percentage of something out of a total)
   - *Averages* (mean value of budget allocations)
   - *Per unit* or *per capita* expenditure (value per person)
2. **Making comparisons** (identifying priority areas and groups)
3. **Analysing trends** (comparing progress over time, adjusted for inflation)

Data on budget expenditure

What governments plan to spend and what the actually spend is often different. Corruption is a leading cause, but inefficient financial management systems, fund diversions and weak oversight can also contribute to the gap.

A variety of tools and methods exist (often called "Follow the Money" approaches) that track data on expenditures, including: government oversight and auditing reports, monitoring the public procurement process and non-governmental oversight and auditing of spending.

Accessing budgetary data

Budgets are official government documents. Generally speaking, they should be available from the website of the treasury or finance ministry, auditor generals offices or anti-corruption agencies. In many countries, however, relevant documents are not made public, and few governments offer appropriate mechanisms public participation in budget processes.

To learn how open your government's budget process is, visit the ***Open Budget Index***, which ranks countries according to the degree to which the public is able to access eight key documents in the budget process (***http://internationalbudget.org/what-we-do/open-budget-survey/***)

Other sources of data relevant to analysing budgets come from international financial institutions, such as the World Bank and International Monetary Fund. NGOs working on corruption can also help you locate budgetary data.

Depending on your region and political context, it may be wise to use a VPN or other anonymising tool to mask your budget data searches.

Case study

Discriminatory budget cuts in Spain

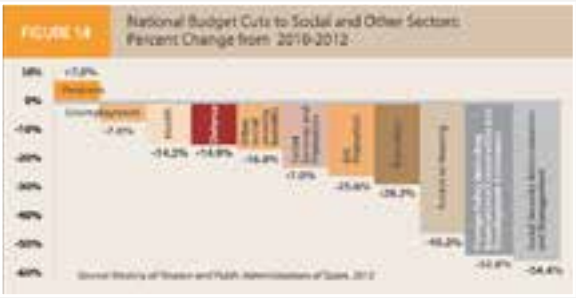
A 2012 study conducted by the Center for Economic and Social Rights analysed Spain's austerity policies from a human rights perspective. Income data showed that a quarter of the population and almost a third of all children were at risk of poverty and social exclusion. There were also vast regional differences.

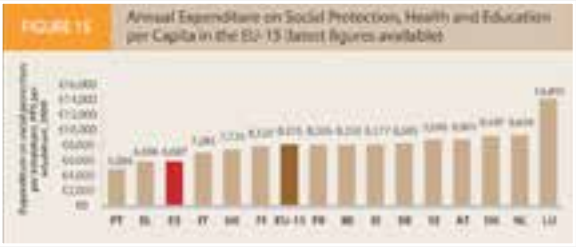
In May 2012, Spain approved the largest budget cuts in its democratic history, totaling €27.3 billion. The rationale was to reduce the public deficit. However, experts warned that these cuts came at the expense of the accessibility and sustainability of basic social services.

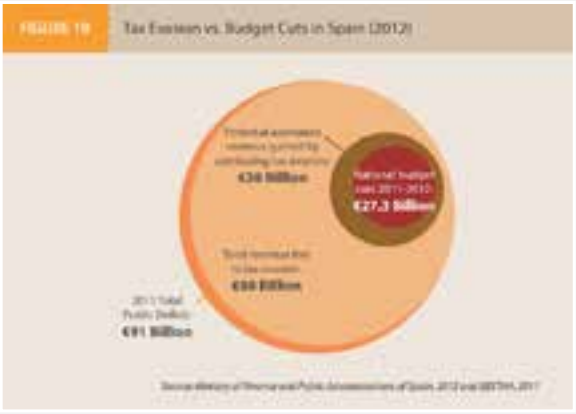
According to widespread misperception, the financial crisis in Spain was thought to be caused by overspending. In fact, Spain was among the lowest spenders in Europe on social protection, health and education.

Spain had a large shadow economy, which caused a significant loss of revenue. Spain's union of tax inspectors calculated that if Spain were to bring the size of its shadow economy in line with European Union standards, it would be able to generate €38 billion, exceeding the total budget cuts for 2012.

Researchers concluded that by focusing almost exclusively on public expenditure cuts, people were being deprived of their fundamental social and economic rights. And without considering alternatives for shrinking the deficit, Spain failed to meet its obligations under the Covenant on Economic, Social and Cultural Rights.







HERE TODAY, GONE TOMORROW: PRESERVING ONLINE VIDEOS AND PHOTOS

Imagine you've found and verified an online video that clearly exposes a violation. The video has the potential to strengthen a case or advocacy campaign you are working on, so you save the link. A few days later, the video is gone...

Online videos and photos can disappear quickly, especially if they are graphic or deal with sensitive human rights issues. From January to June 2015, YouTube took down over 5,700 videos due to government requests alone. This is why you should never treat online platforms as data storage units. Videos can be deleted by the uploader or by the platform for any number of reasons, including a violation of terms of service, user complaints, copyright infringement, account cancellation, or if the platform as a whole shuts down.

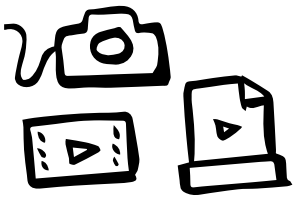
Preserving online content such as photos and videos, including their metadata, will ensure that you can access it in the future, help maintain research integrity, enable you to catalogue the data, and establish a chain of custody for use in legal settings.

How to download

Googling 'download Facebook video' reveals a number of free tools, such as ***VideoVault***, for preserving online photos and videos, also known as "scraping". There are also browser plugins and custom scripts available for downloading large amounts of content. Be aware that scraping content violates the terms and conditions of many sites. It is important to be responsible about how you save, share, and credit/link to the original uploader or content creator.

Beware: there could be malicious trojans in scraping software that, when downloaded, can cause damage to your data and privacy. Look for open-source software that has good reviews, and always download software directly from the developer's site or trusted source.

Your online activities are not anonymous. To avoid surveillance that targets human rights work, consider using a VPN when searching for and downloading sensitive materials.



In depth

Basic preservation steps

1. **Decide what content is most relevant** to your project and which files you want keep. Try to download the largest file size possible and save the files in their original format.
2. **Use a consistent way of organising your photos and videos.** For example, name a folder according to the year, month and day: 2016-03-16. Within the folder, name the files so that they are listed in chronological order: ***00001.AVI***, ***00002.AVI***, etc.
3. **Use a separate text file to document the metadata**. If the footage is violent or graphic, make a clear note of this to caution viewers. The text document should be stored alongside the video or image file. If you're collecting a large amount of content, consider creating a catalogue.
4. **Store in a secure location** such as the home of a trusted ally or in a safe at the office of an established NGO, and back-up on two separate drives, or devices which are kept in different locations. Consider using cloud storage which explicitly focuses on secure data storage, such as SpiderOak or TeamDrive.
5. **Check your files and drives periodically.**

For detailed guides on these steps, visit ***archiveguide.witness.org***

Capturing live streams

Live streams can provide excellent, undoctored evidence as long as they are archived in some form during broadcast. Many platforms automatically archive live streams. If not, you need to anticipate the need and have software installed on your device to capture them.

VPN or "Virtual Private Network", is a technology that creates a secure connection over a public network such as the Internet. Using VPN 'proxies' can help you circumvent Internet filtering. You can learn more about VPNs and proxies in Security in a Box, a guide from Tactical Technology Collective and Frontline Defenders.

ORGANISING YOUR PHOTO AND VIDEO CATALOGUE

When working with large quantities of videos or photos, it's critical to use metadata for effective cataloguing. This will enable researchers to quickly identify the correct files, and save valuable time.

What goes in the collection?

Before building a catalogue, it's helpful to identify the desired outcomes, objectives and uses of the catalogue. This will help ensure that the content you collect and catalogue is relevant.

- **Why does the catalogue exist? What objectives does the catalogue serve?** E.g., advocacy, legal case, media report, public awareness, historic record of a single incident or widespread violations, etc.
- **What is the scope of the catalogue?**
  - What videos or images will you accept? What will you not accept?
  - How will the data be transmitted to you?
  - What sources or formats will you collect? Eyewitness, anonymous, news/media, live stream, etc.?
- **What risks are you willing to take?** Footage that could put you or others at risk or a subpoena?
- **Does your data need to be interoperable with someone else's systems?**

Access: Identify your users and contributors

Understanding who will use and contribute to the catalogue helps you make access and permission decisions. When considering security issues, identify what threats you may face if the catalogue is publicly available or stored on an unsecured platform.

- **Who are the users?** Determine who should have access to your content.
- **How will you handle usage rights?** E.g. Creative commons, attribution to an individual or organisation?
- **Will the catalogue be public or private?** What are the access policies?
- **Who are the contributors?** What level of tech-literacy do they have? What languages do they speak? What file formats do they use?

Establish a workflow and select a platform

- **What is your workflow for ingesting, verifying and cataloguing content?** E.g. eyewitness reports, online form, spreadsheet, printed form, etc.
- **What platforms meet the requirements of your data structure?**
- **What are your technology constraints?** E.g. no internet, limited funds, no technical expertise, etc.
- **Is the platform compatible with your computing environment?**
- **What languages does it need to support?**
- **What are the inherent security risks of the platform?** Are they compatible with your security requirements?
- **Which platforms offer the best security for your data?** Think about the security as the data is being communicated, information 'in motion', and as it is being stored, 'information at rest'.
- **How will you share the information with your end user?**

In depth

Examples of online archives

The ***Chechen Archive*** began collecting photo, video and audio files from the Chechen war, starting in 1994. To make the file information easier to find, use and visualise, they created an extensive catalogue to capture each file's metadata.

The ***People's Archive of Police Violence*** is an online archive of stories, memories and accounts of police violence as experienced or observed by Cleveland, Ohio citizens." See their ***Contributor Terms of Service*** for details on what content they accept.

The ***Syrian Archive*** is an online archive preserving videos documenting human rights violations and war crimes committed by all sides during the ongoing Syrian conflict. As well as building the online database, they support human rights advocates in their documentation efforts.

Structuring your catalogue

Collaboration with end users helps identify the best structure for your catalogue. For example, a human rights lawyers may tell you they need the following metadata: date and location of a recorded crime, type of crime, officer ID and videographer contact information. An advocate, on the other hand, may need to know the name of a victim or where a video was first published.

- **What information does each type of end user need in order utilise the collection?** Consider involving end users in the development process.
- **What is the minimum metadata set?** What elements are required, what is recommended, what is optional?
- **How much time do you have to spend on cataloguing?**
- **Do you need to preserve and catalogue chain of custody because your documentation might be used for criminal prosecutions in the future?**
- **What information needs to be withheld for security purposes?**

Retention: Planning a realistic shelf-life for your catalogue

Knowing the expected period of use for your catalogue enables you to better plan your project, allocate resources and define project success.

- **How long will you maintain the catalogue?** Will you keep certain types of content longer than other types?
- **What happens to the catalogue?** Decide on a succession plan: responsible disposal, handing over of data, etc.
- **Is your retention policy realistic?** E.g. Can you afford to keep what you're planning to for as long as you're planning to? Will you have dedicated staffing to maintain, update and monitor the content for the entire life of the collection?

Examples of platforms

- Google Sheets/Google Forms
- Excel Spreadsheet
- *Filemaker Pro*
- *Martus*
- *Omeka*
- *Corroborator*

Case study

Using video to expose patterns of abuse in forced evictions in Brazil

In 2012, ***WITNESS*** collaborated with activists, lawyers and researchers to amplify and document instances of forced evictions in Rio de Janeiro, and to counter the government's claims that no violations occurred before, during or after forced evictions related to World Cup 2014 and Olympics 2016 construction. They used Google Forms to collect, catalogue, contextualise, and systemise over 100 videos from YouTube, eyewitness sources, activists and media reports. The data was used to create a report on the impact of the widespread evictions.

The graph on the side shows the different types of violations that occurred during evictions.

WITNESS Media Lab offers examples and ideas on how to visualise your catalogue.

Photo: "***Fighting for justice in Brazil***"
Anti-eviction protest in Maua, Brazil 2013 by CAFOD Photo Library on flickr, CC-BY-NC-ND 2.0
https://www.youtube.com/watch?v=2eAIKhFj0m4

AS SEEN FROM ABOVE: SATELLITES AND DRONES

As aerial and satellite information platforms are increasingly used in sectors ranging from forestry and agriculture to city planning and humanitarian aid, the information collected is increasingly becoming available for repurposing in human rights research too.

Aerial photos and drones

Though drones and UAVs (unmanned aerial vehicles) may seem futuristic, they are simply a means to capture views from the air, and are already popular devices for both hobbyists and professionals.

Aerial views offer extensive coverage of the ground that is useful for researching and documenting infrastructure, environmental changes, damage from conflict or natural disasters, and much more.

If you want to launch your own drone there are helpful resources available for ***getting started, deployment***, and for understanding national and local ***regulations*** (be sure to know the rules or you could risk getting fines, confiscation of equipment, or even arrest!).

Photo: AJ+ (Al Jazeera Plus) shares ***video captured by a drone*** flying over Aleppo, Syria in 2015 to clearly show damage to the city after four years of war.

Often human rights investigators will use drone footage found on online platforms like ***Open Aerial Map*** or ***UAViators***. This type of footage, uploaded by individuals or organisations, either anonymous or named, requires the same verification steps as all other online content.

Satellites for human rights

In January 2015, *The New York Times* ran a cover story on a ***Boko Haram*** attack on two towns in northeast Nigeria. In order to show the massive scale of destruction of homes, they used satellite images provided by Amnesty International and Human Rights Watch. It is a powerful example of how satellite imagery can advance human rights research and advocacy.

Satellite imagery is no longer the exclusive domain of governments, and new trends in image supply will make this type of data even more accessible in the future. Both small and large groups are already able to use satellite imagery for daily human rights work.

Benefits of satellite imagery

- **Physical access**
  Satellites enable the circumvention of access restrictions. In armed conflicts, they can be mobilised instead of a researcher.
- **Historical record**
  Satellite images can often allow researchers to travel back in time to study changes on the ground by searching public platforms such as Google Earth.
- **Visuals**
  Satellite images can provide moving and powerful visuals for use in human rights campaigns and advocacy work.
- **Reliability of source and access to metadata**
  In comparison to social media, the original source of a satellite image is usually clear. It often also comes with detailed metadata such as coordinates, date and time, making it a reliable data source.
- **Advanced analysis**
  Satellite sensors detect more information than the human eye, such as infrared or ultraviolet light, which can be used to highlight vegetation and measure vegetation health. These techniques can for instance be used to measure and visualise the impact of oil spills, documenting traces of military movement during armed conflict, or measuring environmental violence as an indicator of genocide.

Limitations of satellite imagery

- **Interpretation**
  Image analysis can be susceptible to misinterpretation, especially if done by analysts or researchers not trained in satellite imagery analysis. For example, disturbed earth can be interpreted as a mass grave or a vegetable garden. Cross corroboration and verification with other information is therefore crucial.
- **Availability**
  Commercial or government satellites do not persistently monitor every location on earth, and there can be glaring gaps in image availability of locations, especially in very remote or politically sensitive areas.
- **Clouds**
  Even if imagery is available at the desired date, cloud coverage can strongly hinder the usability of imagery.
- **Issue areas**
  Satellite imagery is beneficial to a wide variety of human rights issues but can be of very limited use for other areas of concerns. Enforced disappearances, for examples, are difficult or impossible to document through satellite images.

Case study

Exposing a massacre and demolitions in Nigeria with satellite images

Following reports about killings and demolitions in Zaria, Nigeria in December 2015, analysts from Amnesty International reviewed free online satellite images from Google Earth to find supporting evidence for ***large-scale unlawful killings by the Nigerian military***, exposing a crude attempt by the authorities to conceal evidence. Thankfully, in this instance, Google Earth contained up-to-date imagery that could be accessed with the "Historical Imagery" slider in the top left menu. This feature allows you to see how landscapes change over time. In the case of Zaria, it helped identify several demolished areas, including a cemetery and mosque, as well as the appearance of a likely mass grave. The "Save Map" function was used to download images for reporting. All Google Earth images can be used freely for ***non-commercial purposes***. Paying for the ***pro version*** enables downloads of high-resolution images that are better for print.

To aid the visual impact of satellite imagery in online reports, a simple slider can be created allowing a reader to explore temporal changes for themselves. One free and easy tool to create an embeddable slider is ***Juxtapose***.

Example 1: Demolition of Mosque
Before: *https://c2.staticflickr.com/8/7754/26656228073_56a3c2f374_c.jpg*
After: *https://c2.staticflickr.com/8/7236/27261977225_9774490f7e_c.jpg*

Example 2: Mass Grave
Before: *https://c2.staticflickr.com/2/1507/25903760253_1713948455_c.jpg*
After: *https://c2.staticflickr.com/2/1704/25901693544_717cff591d_c.jpg*

See before-after sliders of all locations on the ***Google Earth Blog***.

Example 1: Demolition of Mosque

Example 2: Mass Grave

Sources for sub-meter satellite imagery

- **Google Earth**: Can be a highly useful tool to source free satellite imagery. However, it has occasionally reduced resolution in some areas as a result of political pressure.
- **TerraServer**: A provider of DigitalGlobe satellite imagery. It can be used for free in support of basic research, and often provides more current imagery than Google. You will have to pay a small fee for downloading the actual images. Additionally, you must pay for a license if you'd like to use the imagery publicly.
- **Commercial image providers**: The main providers for commercially available very high-resolution satellite imagery are ***DigitalGlobe***, ***Airbus***, and ***Urthecast*** (***Deimos-2***). The resolution for these satellites ranges from 0.3-0.75 meters (roughly the smallest size of an object/subject that could be identified in an image). The minimum order size for purchasing imagery is 25 km2. Prices for this size start at approximately $175.
  - **Microsatellites**: A big advantage of companies such as ***Planet Labs*** or ***TerraBella*** is that they will eventually provide a full constellation of satellites, i.e. they are creating a system of constant monitoring from space. The companies mentioned above can only provide snapshots of certain areas on earth. Be on the lookout for ***satellite video***, which will have big implications for human rights research. A downside of these microsatellites is that most of them currently have a lower resolution, i.e. fewer details are visible in the imagery.

Responsible data considerations

REAL LIFE SAFETY RISKS

Physical safety and security in the context of digital data requires a different lens and approach. For traditional documentation collection, human rights professionals often follow intuitive best practices to protect themselves and the individuals they interview: meeting secure locations, concealing identifying information confidential, and avoiding going directly into harm's way.

Physical security for new data methods are, for many, not yet as intuitive because there is often no personal relationship between the researcher and the individuals referred to in the data.

Physical and digital security are closely linked. The purpose of digital security is not only to protect the data but also to protect the individuals who contributed to or were featured in them.

If a hacker compromises a server or a checkpoint official confiscates a hard drive, the names, faces, and/or information of a large number of victims or other vulnerable individuals are compromised. Your obligation is to take steps to protect yourself and others from additional harm.

The people involved in data fall into three sometimes overlapping categories: those who capture the data, those who share it online or on an external hard drive, and those whose information is contained in the data.

Often people do not realise that a video they are posting online contains sensitive information. If you, as a human rights professional, decide to use a video for a campaign, the source and creator may be targeted for being affiliated with your group, even if they've never heard of you.

Those featured in the materials may not have consented, or even been aware that they were being filmed. If you publicly reveal their faces or names in the context of a human rights violation, you may inadvertently cause them to be harassed or further abused.

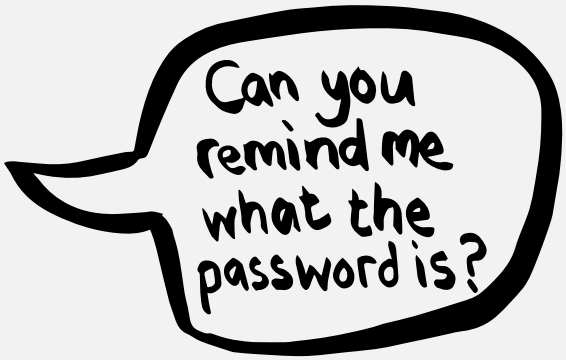
In the event that you publicise the data, or there is an external breach, the source and featured individuals may be at increased risk of physical harm even if their information was already posted online. Worse, it may be difficult to inform any of the at-risk individuals of the heightened risk, as the researcher is unlikely to have contact information or other means of reaching those identified. Therefore, avoiding the release of data in the first place is critical.

Case study

Social engineering: Who are you really talking to online?

Human rights activism in Iran provides a striking example of the connection between digital and physical security. Recently, Iranian authorities ***have used social engineering*** to target human rights defenders through communications via email, Facebook, and LinkedIn. Social engineering is the act of using personal facts and psychological manipulation to convince an individual that it is safe to break security protocols and divulge private information. ***A government agent could create a social media profile*** using the name of a known activist and initiate conversations that trick other activists into revealing sensitive information.

The Islamic Republic has used this tactic quite successfully, leading to the detention and interrogation of prominent human rights defenders Even ***US government officials have fallen victim*** to social engineering, unintentionally divulging information to the Iranian government. Social engineering is often more effective than traditional hacking because it does not rely on any technical skills and preys on the vulnerabilities of individuals rather than digital systems.



To avoid being attacked through social engineering techniques, be vigilant in all your communications. Ask clarifying or personal questions to ensure that the person you are speaking to is who they say they are and discontinue the communication if you have any doubts as to the identity of the account holder. Also avoid sharing sensitive information over social media or email without additional security measures in place.

RESPONSIBLE DATA

Responsible data is:
The duty to ensure people's rights to consent, privacy, security and ownership around the information processes of collection, analysis, storage, presentation and reuse of data, while respecting the values of transparency and openness.

Responsible Data Forum, working definition, September 2014

To stay up to date on responsible data issues, see ***https://responsibledata.io*** and sign up to the mailing list at ***https://engn.it/rdmailinglist***

Being thoughtful and proactive about how you work with digital data is essential to ensure the safety and well-being of people you work with. The umbrella of ethical issues that might arise can be talked about as **responsible data challenges**. There is no set policy that will work for all situations, but we encourage you to think of the following issues within your context:

- **Informed consent**
  Especially within the context of human rights violations, getting informed consent can mean much more than you first think. Make sure that people you're getting consent from are giving it voluntarily, that they fully understand what they are consenting to, and that they're in a position to make these kinds of decisions.
- **Data minimisation**
  In the interests of privacy and lowering risk, always attempt to collect the minimum amount of data necessary for your specific purpose.
- **Personal data**
  Whenever possible, if you have names, dates of birth or other personally identifiable information, try to de-identify it or delete what you don't need.
- **Sensitive data**
  Be conscious of where sensitive data on race, ethnicity or sexual orientation is stored, whether on your own servers or in a foreign country (eg. by using a Google service) in case it could expose individuals to risk.
- **Implicit bias**
  Data gathered from other sources may have hidden biases. Think carefully about the human decisions that were made during analysis.

Information sourced from social media

Many organisations seem to think that information 'in the public realm' is free to use it for any purpose. However, even if a person agrees for their information to 'go public' in a certain way, it does not mean the information can ethically be appropriated for any use.

When a person posts about a human rights violation on Twitter, it does not mean they automatically consent to becoming the focal point of a human rights campaign with global reach. You need to contact people directly and help them understand the possible consequences (positive and negative) of attracting more attention, so they can make an informed decision.

Protecting people's identity on social media

Social media posts which document violations may contain video and/or images of people who had no idea they were being filmed. Whether you are the one publishing this data or an organization that would like to re-release it you must protect individuals as much as possible.

There are a number of helpful tools to blur faces in videos and photographs. For example ***ObscuraCam*** is an android phone app that pixelates or redacts faces. It works either while you are recording and when you load something to it from your library. YouTube has also recently added a ***blurring feature*** that allows you to obscure faces when you load a video.

ARE YOUR DIGITAL TOOLS SAFE?

When it comes to making backups, secure file storage, and safe destruction of data, we recommend consulting Security in a Box by Front Line Defenders and Tactical Technology Collective.

Digital security is not limited to using secure tools and software, but is embedded in general organisational habits and practices.

When we talk about digital security, we are talking about taking steps to protect data both while in transit (email, chats, and SMS) and while at rest (a file saved on our hard drive).

We take these steps because in collecting, keeping and sharing sensitive data, especially data related to human rights, there are implications for ethics and safety.

We have a responsibility to protect the privacy of those whose names and identities are represented in our data, and who trust us with their information.

New tools and services make various claims about the security of their products. How do we know that it's true? These are some of the questions we should be asking when considering new software. You are unlikely to be able to answer all of them, but do the best you can.

Before adopting new software, ask yourself...

1. **What does "secure" really mean?**
Saying software is "secure" could mean a variety of things. Read the fine print of any software description. Many tools that claim to be "secure", including Dropbox, Google, and Skype, encrypt your data as it travels between your computers and their servers – but the data is not protected from being accessed by the companies themselves. Always ask how a service protects your data. Some tools and services use encryption in ways that protect the data even from the service providers themselves. It is important to understand this distinction, and to consider whether it is important to your project to keep data hidden from service providers.

2. **Has the software been independently audited?**
Don't just trust software developers when they promise to offer security, especially when software is brand new. Even with the best intentions, implementing encryption correctly is hard. Only use software that has been independently audited by someone you trust. Preferably the code of the software should be "open source" and available for public scrutiny. And even then, you can't be sure whether clean audits have missed vulnerabilities. Maintain healthy skepticism always.

3. **Who owns the data–and for how long?**
When you use an online service to communicate, create or archive information, you should question who "owns" the data, especially if it is stored on the servers of a service provider like Facebook, Twitter or Dropbox. The answers are typically buried in a company's Terms of Service. Whatever service you use to archive or share sensitive data, you should research their policies on sharing data with other companies or governments, and what you can expect to happen to the data if you delete your account or the company gets sold or closes down.

4. **Does the software offer protection from the particular threats you face?**
Understanding your own threats is important in assessing whether an application or service will adequately protect you. Rather than adopting a tool because you heard you should, try building a "threat model" to establish *who* may try to access your data, and *how* (see The Surveillance Self Defense Guide's ***Introduction to Threat Modeling***). The answers to these questions will help you establish which tools are the most appropriate. To learn more about carrying out a context analysis, see the Explore section of the ***Holistic Security Manual***.

5. **Will your use of the software or service expose you?**
In some places using encryption is illegal. In others it is not, but will still draw suspicion from law enforcement or intelligence agencies. This may lead them to begin monitoring a person's activity simply because they have begun using security or encryption software. It is important to understand your risk; in some contexts, *not* using encryption may be safer, even though the data is stored and sent unprotected (which is also risky!). It may allow you to keep a lower profile and attract less attention. Knowing your context is essential to making this kind of assessment.

6. **Is it practical for you to adopt this service or software?**
Some software or services may fit well into existing processes you're using to collect, store, analyze, and share data. Others may demand a major shift to new processes. Even for software that promises security, practicality is an important question to ask. If it's too hard to use or is not likely to be a sustainable practice, attempting to adopt it may end up being a waste of valuable resources in time, money, and energy.

7. **Are you adopting this application for the right reasons?**
There are many reasons we may be interested in a new service or tool: it is designed beautifully, it has an interesting name, it promises a novel service, it is growing in popularity, etc. Take a hard, critical look at the reasons you're considering adopting a tool and make sure those feel like the *right* reasons, given your ethical and security needs and obligations.

It is very rare that any software or service meets all of these needs. The goal is not to find a unicorn that satisfies *every* requirement, but to make informed, intentional decisions about the software we adopt. After reading this advice, do you feel differently about any tool or service that you have recently adopted? Does it remain the right choice for you?

In depth

Upfront about vulnerabilities.

Some developers publish "threat models" and known vulnerabilities on their website for the sake of transparency. For good examples, see documentation for ***Scramble.io*** (secure email software) and ***Cryptocat*** (secure chat software). More companies now also release transparency reports that highlight things like official requests for user information and legal takedown notices. For example, see ***Twitter's Transparency Report***. These reports are voluntary and difficult to verify, still they may offer context around the safety of services.

Example

Should Jon use WhatsApp for human rights work?

Jon is a human rights activist in Kampala, Uganda, and is collecting reports of violence against LGBTI people. Homosexuality is illegal in Uganda, and Jon's group is routinely monitored by police. Jon just bought his first smartphone, and is considering using the chat application WhatsApp. Many of his friends use it, and he has heard that it's secure.

Here's what Jon decided...

WhatsApp implemented widely-trusted encryption by Signal in 2016, which is open source and verifiable. But WhatsApp itself is not open source and there are no public audits of the code, so Jon can't be entirely sure of their security. Also, WhatsApp messages are encrypted in transit, but not on the device, so a stolen phone could be vulnerable. Jon does not outrule using WhatsApp, but asks a colleague, and decides to research other messaging applications, including ***OpenEvsys*** and ***Martus*** which are developed for human rights purposes.

SECONDARY TRAUMA AND PTSD

Human rights professionals are committed to helping others. As such, well-being and security are often considered luxuries or selfish objectives, particularly when operating in conflict zones. It is essential that both organisations and individuals work to counter this misguided belief. The recognition of the importance of security and well-being fosters resilience and agility, improves the management and mobilisation of resources, and enables preparation for the risks inherent to human rights work.

Security relates to physical safety, health, finances, discrimination, privacy, etc. Security threats vary from person to person and group to group. For some their religion or sexual orientation may pose the greatest safety risk. Understanding these concerns from the perspective of everyone involved in your work is the first step in fostering a productive and safe environment.

Conducting and maintaining the following analysis will help you create strategies, plans, and shared agreements to promote security and well-being:

- Explore the political, economic, social, technological, legal, and environmental developments inherent to your work.
- Identify and analyse concrete safety threats and take the necessary steps to prevent or respond to them.
- Identify your allies and opponents, their interests and their capacity to empower or act against you.
- Map and categorise private information and data and take measures to protect it from loss or damage.

Offsetting secondary trauma from your data

The unprecedented volume of audiovisual content available to today's human rights researchers has considerable benefits and considerable risks. Research by ***Eyewitness Media Hub*** shows that 82% of human rights researchers see distressing imagery sitting at their desks several times a month.

Exposure to so much primary trauma can lead to secondary trauma which can, in turn, lead to post-traumatic stress disorder (PTSD). Organisations, managers, and researchers must recognise and mitigate these risks. Eyewitness Media Hub lists several triggers that can be distressing for individuals, many of which are common in digital data sources, including:

- **Surprise**: the individual does not expect to watch at a violent video.
- **Repeat exposure**: the individual has to watch a violent video repeatedly.
- **Personal association**: the content reminds the researcher of a personal situation or a personal connection.
- **Audio of human suffering**: hearing the sound of violence made a video more distressing.
- **Feelings of guilt**: human rights researchers report regular guilt for feeling trauma over violence that is being inflicted upon someone else.

Signs of secondary trauma

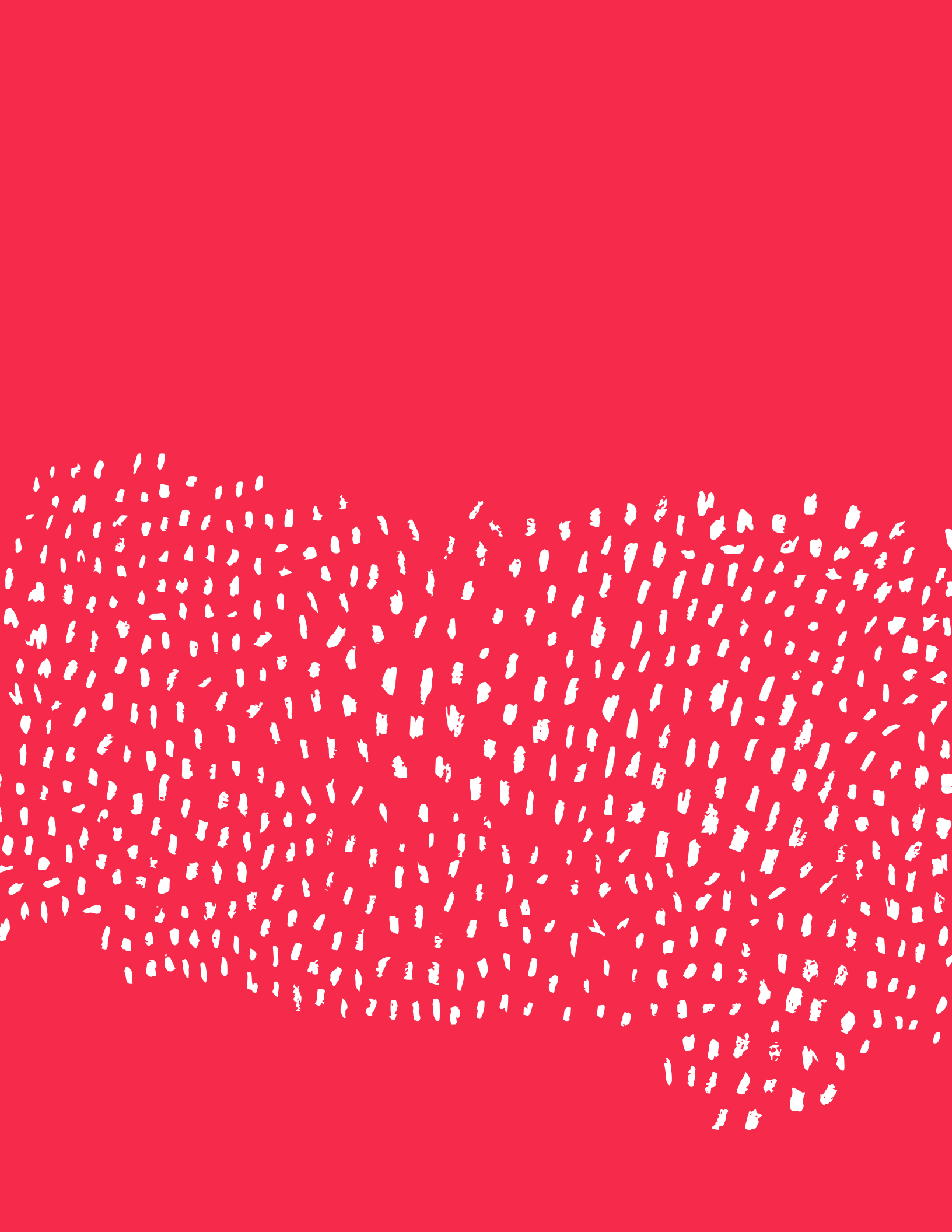
- Difficulty managing your emotions.
- Difficulty accepting or feeling okay about yourself.
- Difficulty making good decisions.
- Problems managing boundaries. E.g., taking on too much responsibility, having difficulty leaving work at the end of the day, trying to step in and control other's lives.
- Relationship problems.
- Physical problems such as aches and pains, illnesses, accidents.
- Difficulty feeling connected to what's going on around and within you.
- Loss of meaning and hope.
- Increased substance abuse, alcohol in particular.
- Binge-eating.
- Isolating oneself from friends and colleagues.

Individuals experiencing secondary trauma may experience none, some, or all of these symptoms. Any sudden behaviour shifts should be investigated.

Immediate steps to avoid secondary trauma

1. Pace the frequency with which staff is exposed to traumatic content.
2. Eliminate needless repeat exposure.
3. Review sorting and tagging procedures to reduce unnecessary viewing.
4. Experiment with different ways of viewing. Some find concentrating on certain details, for instance clothes, and avoiding others (such as faces) can help create emotional distance.
5. Adjust the viewing environment: reduce the size of the image or video or adjust the screen's brightness/resolution.
6. Turn sound off when possible.
7. Take frequent screen breaks. Look at something pleasing, take a walk or stretch.
8. When emailing graphic content, write a content warning into the subject line.
9. Clearly label all files when archiving, so that people are not accidentally exposed to graphic content.
10. Craft your own self-care plan. Research shows that highly resilient individuals are more likely to exercise regularly, maintain outside interests and enthusiasms, and invest time in their social connections when challenged by trauma-related stress.
11. Set up peer support networks within organisations to talk about the horrible content you have had to encounter.

PTSD arising from secondary trauma is professionally treatable. If you think you or a colleague may be suffering from PTSD, seek professional help immediately.



Where to go from here

HOW TO FRAME YOUR RESEARCH

Before diving into analysis, think carefully about what your data represents and what valid conclusions you can draw from it. This section helps establish the representativeness of a dataset and pinpoint uncertainties to avoid inappropriate or incorrect conclusions.

Establishing what type of sample your data is

There are three types of samples. The type of sample you have depends on how many units of analysis you have and how they were selected.

- **Complete sample**: an entire population. Other terms for a complete sample are a "census" or a "complete enumeration".
- **Probability sample**: units in the population were selected through a probabilistic mechanism. This mechanism ensures that the probability of selection for every unit in the population is known.

Convenience sample: units were selected by a non-probabilistic mechanism. For each unit in the population, the probability of being selected into the sample is unknown, while it is very high for some units and very low or zero for other units. In human rights research, convenience samples are the most likely type of available data samples.

Determining what conclusions you can draw

The process of learning and drawing conclusions from data is called "inference." There are two types of inference:

- **Statistical inference** draws conclusions *about the population*, based on the sample. These conclusions are valid only if you have a complete or probability sample.
- **Description** draws conclusions *about the sample* you have.

Making inferences from complete samples

Complete samples are special in that they are identical to the population, so whatever you learn about the sample, you also learn about the population.

Making inferences from probability samples

Probability samples are representative of the population if the underlying sample probabilities are accounted for. Because each unit in the population has a known probability of being selected into the sample, on average, we can approximate the same values and distributions in our sample that exist in the population (minus a probability sampling error). We do not consider all units in the population, but only a representative sample thereof, so we call these conclusions "estimates".

In depth

Population

In statistics, population refers to all the units of analysis you want to study.

Example 1
You want to learn how many refugees entering country X by boat are being denied access to health care. The population is all refugees who arrived in country X by boat and were denied access to healthcare.

Example 2
You are examining whether commercial flower plants have caused water pollution in a designated region. Here, X is all flower plants in that region.

Convenience samples

Convenience samples are not representative of the population: they are biased in unknown ways. All convenience samples contain selection bias. In statistical terms, this means we do not fully know how well our sample represents the population. We do not know which parts of the population we are not seeing, how big of a proportion they constitute of the population, or how our sample is different.

Making inferences from convenience samples
Convenience samples are not representative of the population. They are biased in unknown ways, as we do not know which population units were left out, or which units were less or more likely to be selected. We can not learn about our population, and we must not overstate the conclusions that we can draw from the sample.

Conclusions you can draw from convenience samples

- The number of people you were able to document whose right to health care was violated.
- Summarise and describe the individuals concerned and medical issues they suffered.
- Create graphs to visualise the violations you documented. Be sure to clearly identify these graphs as summaries of "reported violations" or "documented violations." (See ***this HRDAG report*** for an example of how to describe convenience sample data and how to annotate graphs of such data.)
- Describe the state practices you've identified that lead to a denial of healthcare access.
- State where and at what times you have identified these practices. If you have documented these practices across several healthcare facilities throughout country X and/or repeatedly over time, you can point out this spatial and/or temporal spread.

- Point out that the absence of additional reports of such state practices does not mean that those practices did not occur. You may have simply been unable to identify more of them. To support this statement you could list the facilities you were unable to visit in an effort to collect further evidence. Missing evidence may not indicate a lack of violations, but simply be the result of not being able to research in that area.
- Conclude that state practices are an issue to be taken seriously.

Conclusions you cannot make

- Number of refugees denied access to healthcare. Because this is a convenience sample, you do not know how many refugees you were unable to document.
- Deduce which state practices are most and/or least common.
- Deduce spatial trends by stating whether state practices are more prevalent in one area compared to another.
- Deduce temporal trends by stating whether state practices occurred more frequently during one time period compared to another, and/or whether state practices increased, decreased or stagnated over time.

If you treat information in your convenience sample as what it is -evidence of human rights violations you were able to document- you will have strong grounds for justifying your human rights claims in public.

Documenting your research process

- Clearly outlining the methodology of how you reached a conclusion makes your findings accessible and transparent, adding legitimacy to your claims.
- Internal documentation and archiving saves you time and money, while reducing the risk of information loss.
- Documenting successful practices furthers your knowledge within and beyond your community. This is especially relevant with digital data sources, where the knowledge base is still developing.
- Take note of all the decisions you made and map out the data sources you used. Describe how you obtained each data source and how you worked with it.
- A step-by-step summary of your research methodology should enable others to follow your steps, repeat your findings, and draw comparable conclusions.

Sample questions your methodology should attempt to answer

- **What is the geographical focus of your study?**
- **What is the temporal focus of your study?**
- **How did you measure X?**

  a. Data source #1

  i. Where and when downloaded?
  ii. Who was the original creator/provider? Is this source credible? How did you verify this data?
  iii. Publically available?
  iv. Did you systematically organize this data in some way? For example, by hour, or by day?
  v. Are there any restrictions, concerns, or challenges with this data?
  vi. How did you deal with missing information?
  vii. Search terms used, hashtags followed? In which languages?

- **What methods did you use to obtain the results from the collected data?**

  b. Data processing and management

  i. Process of translation, if applicable?
  ii. Software used?
  iii. Process of combining/linking different data sources?

  c. Data analysis

  i. Software/programming languages used?
  ii. Statistical methods employed?
  iii. Procedure for dealing with missing information?
  iv. Potential types of robustness checks and sensitivity analysis?

Before you call in the experts... prepare your questions

Researchers have more easy access than ever before to data. But many still lack the abilities or confidence to work with large data sets. In the already cash-strapped and overworked field of human rights, this puts data experts under particular strain, as they struggle to meet the many demands of their colleagues. The more educated and empowered you are to do your work, the more stress and time you will save yourself and your colleagues.

That said, sometimes you need to call in the experts. Before you do, prepare for your conversation by considering the following questions:

- **What is your end goal?** Being explicit about your aims helps the expert understand your request.
- **What does good support look like?** Are you seeking assistance understanding bias in a specific dataset? Support analysing a data stream that is new to you? Confirmation that you are on the right track?
- **Do you need specific or ongoing assistance?** Be realistic and only ask for the amount of assistance you can afford.
- **Why this expert?** Have they worked on similar cases before? Did someone recommend them? Have you considered alternatives?

Do your homework

Look for projects that have done similar things to what you want to do. A good place to start is with the case studies you'll find in this set of resources. What specifically do you like about those projects, with regards to their use of new or emerging data sources? Are they using new types of data in a way you hadn't seen before, or do they manage to draw new conclusions based on a new type of analysis?

Look for examples of similar projects that you like, which use the new types of data in the particular way that you'd like to. Once you find them, write down the date of when they were created, and pick out any elements you find particularly inspiring, or are particularly relevant for your work. Once you've got a list of projects, reach out to some of the researchers involved in them and see if they have time for a quick call with you. Be realistic: though a particular project might seem very simple from the outside, it might have involved more work behind the scenes than you might realise.

Stay critical

Though you have questions, you still have the most expertise on your specific context. If someone tries to tell you that something is either *always* or *never* right, and you disagree, trust your instinct.

Especially when money becomes involved, get multiple opinions on what would be the best option for you. Software vendors, whether motivated by profit or limited by a focus on their own products, may not give you the best recommendation. Consider more secure and generally less expensive Free and Open Source Software (FOSS) tools wherever possible.

CONCLUSION

There are a number of organisations and individuals who are using these techniques to push the boundaries of human rights documentation, in new and exciting ways. Many of them have been strong sources of inspiration for this guide, and for our collective work as a whole, providing us with just a glimpse of what exists at the frontiers of data use in human rights documentation.

Groups like ***Forensic Architecture***, who have done projects which use innovative data modelling techniques to project what happened in the past based on limited pieces of information that remain after the fact. Like 3D modelling the environment around a ***boat in the Mediterranean***, which was left to drift after no country took responsibility for it. They produced a report which formed the basis for a number of ongoing legal petitions filed against NATO member states.

Or ***Bellingcat***, a network of citizen investigative journalists who use publicly available information, including social media data and more generally, Open Source Intelligence (OSINT), to investigate human rights violations and more. Bellingcat founder Elliot Higgins' work in 2013 investigating the ***chemical weapons attack on Ghouta, Syria***, in August 2013 helped prove that the perpetrator of the attack was almost certainly the regime of Bashar al-Assad. They have produced a fantastic set of online resources, including ***case studies*** and tutorials.

Statisticians at the Human Rights Data Analysis Group have used innovative techniques to accurately estimate civilian casualties in war in a number of countries, pioneering a technique called Multiple Systems Estimation. Among other successes, their founder ***Patrick Ball has testified in the Guatemalan Supreme Court*** against former Head of State General José Efraín Ríos Montt, who was then found guilty of genocide and crimes against humanity.

Regarding verification, organisations and individuals are making their processes more transparent. Tools like ***Checkdesk***, ***Github*** and ***Jupyter Notebooks*** enable people to publicly document their processes and conclusions, increasing credibility, and allowing anybody to follow and learn from them. It also opens research up to a new degree of scrutiny, beyond that of more traditional human rights documentation.

The future is ours and the time is now

Between your human rights expertise and the information in this guide, you have everything you need to assess the value of different data for your work, and to begin using them.

The considerations, assessments, and uses outlined are rooted in practices with which you're already very familiar–like those around ethics, security, and verification.

Developing the ability to critically assess and understand what new tools and platforms allow you to do and how remains much more important than focusing on specific tools. Being aware of what those tools do, the biases within the data you receive and work with, and, above all, engaging with the data responsibly, are as important as ever.

At the time of writing, projects like those discussed above are far from being the norm. Generally speaking, human rights researchers are still using the same techniques that they have been for decades.

Available data will continue to increase and become more accessible. Its costs, as well as the costs of technologies will continue to decrease. Though the learning curve may be steep at times, engaging these techniques is central to the human rights work of today and tomorrow.

RESOURCES AND FURTHER READING

Make the most of the organisations and resources working in the field. Organisations like ***The Engine Room*** provide technical support for organisations wanting to use tech and data more strategically in their work, and the ***Responsible Data community*** is a good place to go for advice on ethical, legal or privacy challenges arising from the use of data in new and different contexts. Groups like ***Tactical Technology Collective*** provide guidance on using information for advocacy, like this ***Visualising Information for Advocacy*** book. ***School of Data*** is home to a strong community working to help civil society groups and journalists use data to tell stories. Their online courses cover everything from cleaning data to analysing it to visualising it. If you want support on setting up and using databases in your documentation, take a look at ***Benetech***, or ***HURIDOCS***.

For links to more detailed guides and resources visit: ***https://engn.it/datnav***

Social media data

Citizen Evidence Lab ***https://citizenevidence.org***

First Draft News ***https://firstdraftnews.com***

Citizen Media Research and Verification: An Analytical Framework for Human Rights Practitioners
http://www.cghr.polis.cam.ac.uk/publications/cghr-practitioner-papers-series/paper-1

WITNESS
https://lab.witness.org

Ethical guidelines for using eyewitness footage
https://lab.witness.org/announcing-witness-ethical-guidelines-for-using-eyewitness-footage-in-human-rights/

Budget data for human rights

Center for Economic and Social Rights (2015), Defending Dignity: a manual for national human rights institutions on monitoring economic, social and cultural rights

International Budget Partnership (2010), A Guide to Tax Work for NGOs

Christian Aid (2011), Tax Justice Advocacy: A Toolkit for Civil Society

Fundar and International Budget Partnership (2004), Dignity Counts: a guide to budget analysis to advance human rights

International Budget Partnership (2014), Article 2 and Governments' Budgets

OHCHR (2010), Human Rights in Budget Monitoring, Analysis and Advocacy: Training Guide

International Budget Partnership (2008), Our Money, Our Responsibility: A Citizens' Guide to Monitoring Government Expenditures

Hakikazi Catalyst (2006), Follow the Money: A Resource Book for Trainers on Public Expenditure Tracking in Tanzania International Budget Partnership

Open Knowledge Foundation, Open Spending Handbook, available at: ***http://community.openspending.org/research/handbook***/

Here today, gone tomorrow: Preserving online videos and photos

To scrape or no not to scrape ***http://www.storybench.org/to-scrape-or-not-to-scrape-the-technical-and-ethical-challenges-of-collecting-data-off-the-web***/

Best practices in scraping: from ethics to techniques ***https://goo.gl/hovkap***

Chilling Effects ***https://lumendatabase.org–database*** that collects and analyses legal complaints and requests for removal of online materials including videos; the Google Transparency Report openly cooperates with this documentation ***https://www.google.com/transparencyreport/removals/copyright/faq/#chilling_effects***

Takedown Project ***http://takedownproject.org*** – partner of Chilling Effects project; effort to mobilize the research community to explore how notice and takedown procedures operate in the U.S., Europe, and other countries, and how these procedures resolve conflicts between copyright and freedom of expression

Electronic Frontier Foundation ***https://www.eff.org/issues/intellectual-property/guide-to-youtube-removals*** has a useful guide and diagrams on YouTube's takedown policies and how to challenge them

Learn more about how satellites are used to expose human rights abuses ***https://www.theguardian.com/global-development-professionals-network/2016/apr/04/how-satellites-are-being-used-to-expose-human-rights-abuses***

Introductory guide to satellites and satellite image analysis ***http://landscape.satsummit.io***

Create a before/after slider: select satellite images and host them online, for example on flickr. Then copy and past the two links from flickr (link must end in .jpg) into the juxtapose tool ***https://juxtapose.knightlab.com/#create-new***

Get inspiration from the Seeing From Above issue of the Exposing the Invisible program at Tactical Technology Collective, which highlights cases of aerial imagery in different contexts through interviews, commentary, and how-tos ***https://exposingtheinvisible.org***

Advanced guides on using satellite imagery in human rights work: Monitoring Border Conflicts with Satellite Imagery: A Handbook for Practitioners ***http://www.aaas.org/report/monitoring-border-conflicts-satellite-imagery-handbook-practitioners***, Satellite Imagery Interpretation Guide: Intentional Burning of Tukuls ***http://hhi.harvard.edu/publications/satellite-imagery-interpretation-guide-intentional-burning-tukuls***

Digital security checklist to determine who should have access to your content ***https://docs.google.com/document/d/17gRphFoh7PlUrmeQNu5ADhNfty-2sEV-CLQN4Ed1hak/edit***

As seen from above: Satellites and drones

Human rights relevant drone footage ***https://www.youtube.com/playlist?list=PLRK6YeiwsEtmkkCikDM8mKSHuo9VDiE_0***

New Technologies for Property Rights, Human Rights, and Global Development ***http://drones.newamerica.org/primer/DronesAndAerialObservation.pdf***

iRevolutions ***https://irevolutions.org/category/dronesuavs/***

Unmanned Aerial Vehicles in Humanitarian Response ***https://docs.unocha.org/sites/dms/Documents/Unmanned%20Aerial%20Vehicles%20in%20Humanitarian%20Response%20OCHA%20July%202014.pdf***

Real life safety risks

New Tactics for Human Rights Activism has compiled a helpful list of considerations and tools for the protection and self-preservation of human rights professionals ***https://www.newtactics.org/conversation/staying-safe-security-resources-human-rights-defenders***

Privacy, Responsibility, and Human Rights Activism ***https://www.newtactics.org/conversation/staying-safe-security-resources-human-rights-defenders***

Towards Holistic Security for Rights Activists ***https://holistic-security.tacticaltech.org***

UNESCO's Protecting Journalism Sources in the Digital Age contains many tips that also apply to human rights researchers who want to protect their contacts ***http://www.unesco.org/new/fileadmin/MULTIMEDIA/HQ/CI/CI/images/Themes/Freedom_of_expression/safety_of_journalists/Protecting_Journalism_Sources_in_Digital_Age_UNESCO_Flye.pdf***

Where to go from here

Are your digital tools safe?

Tactical Technology's "Security in a Box"

Backing up your software
https://securityinabox.org/en/guide/backup

Secure file storage
https://securityinabox.org/en/guide/secure-file-storage

Destroying information
https://securityinabox.org/en/guide/destroy-sensitive-information

Secure app scorecard
https://www.eff.org/secure-messaging-scorecard

The Responsible Data Forum's Handbook of the Modern Development Specialist
https://responsibledata.io/resources/handbook/chapters/chapter-01-designing-a-project.html

Setting Up the Data Infrastructure
https://responsibledata.io/resources/handbook/chapters/chapter-02-managing-data.html

An Introduction to Threat Modeling
https://ssd.eff.org/en/module/introduction-threat-modeling

Threat Modeling for Campaigners and Activists
http://www.mobilisationlab.org/threat-modeling-for-campaigners-and-activists

EFF Surveillance Self-Defense Guide
https://ssd.eff.org/

Holistic Security Manual
https://holistic-security.tacticaltech.org

How to frame your research

Human Rights Data Analysis Group (HRDAG): Core Concepts *https://hrdag.org/coreconcepts*/

Responsible Data Forum: Recognising uncertainty in statistics (Brian Root, HRW)
https://responsibledata.io/reflection-stories/uncertainty-statistics/

Kelly Greenhill: Nigeria's Countless Casualties Foreign Affairs *https://www.foreignaffairs.com/articles/africa/2015-02-09/nigerias-countless-casualties*,Tufts University
http://as.tufts.edu/politicalscience/sites/all/themes/asbase/assets/documents/newsEvents/2015febForeignAffairsGreenhill.pdf

Selected HRDAG publications on selection bias
https://hrdag.org/publications/big-data-selection-bias-and-the-statistical-patterns-of-mortality-in-conflict/

Reports that document the underlying research methodology as one or several sections in a report:

https://hrdag.org/wp-content/uploads/2013/02/Gohdes_Convenience-Samples.pdf

https://hrdag.org/wp-content/uploads/2013/02/results-paper.pdf

https://targetedthreats.net/

https://hrdag.org/wp-content/uploads/2015/07/HRDAG-SY-UpdatedReportAug2014.pdf

https://hrdag.org/wp-content/uploads/2013/02/uv-estimates-paper_2012-11.pdf

https://hrdag.org/wp-content/uploads/2013/02/Benetech-TRC-descriptives-final.pdf

https://hrdag.org/wp-content/uploads/2013/02/Benetech-Report-to-CAVR.pdf

https://hrdag.org/wp-content/uploads/2013/02/State-Violence-in-Chad.pdf

Before you call in the experts... prepare your questions

Video: WITNESS *https://witness.org*/

Statistical methods:
the Human Rights Data Analysis Group (HRDAG) *https://hrdag.org*

Integrating + understanding data and technology strategically:
The Engine Room *https://theengineroom.org*

Secure data collection + software needs // Digital security:
Benetech *http://www.benetech.org*

Document management:
Huridocs *https://www.huridocs.org*

Eyewitness media:
the Eyewitness Media Hub
http://www.eyewitnessmediahub.com

Tactical Technology Collective
https://tacticaltech.org

Networks/communities:
New Tactics in Human Rights
https://www.newtactics.org

Open Government Partnership
http://www.opengovpartnership.org